InSiDE

Innovatives Supercomputing in Deutschland

No 25-1 · Spring 2025

JSC Goes Exascale

The center's JUPITER system is coming online. When fully operational, it will become Europe's first exascale supercomputer.

Energy Efficiency in the AI Era

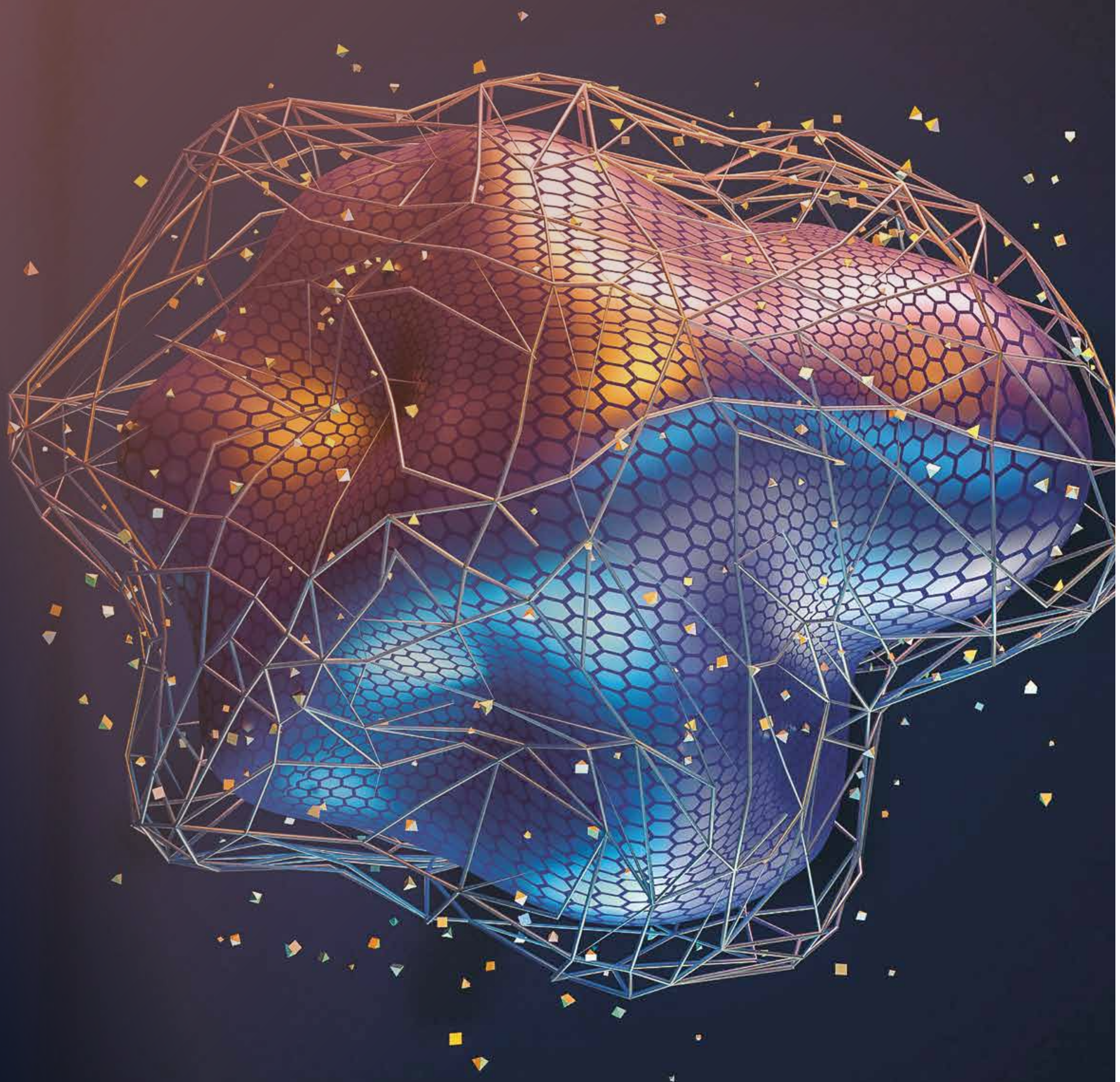
GCS centers are embracing infrastructure, hardware, and software innovation to improve energy efficiency of their systems.

AI Factories Come Online

All three GCS centers are playing prominent roles in new EuroHPC Joint Undertaking awards aimed at bolstering European AI capabilities.

GCS
Gauss Centre for Supercomputing

Artificial intelligence is reshaping countless corners of our society. From improving the efficiency of organizational and administrative tasks to expediting development of new medications, these methods show great promise. They also come with myriad legal, technological, and ethical challenges. The Gauss Centre for Supercomputing supported early research into machine learning applications, and as the AI revolution continues, GCS is supporting both cutting-edge AI innovation as well as helping build out a robust foundation for the technology's responsible, effective growth. This issue highlights just a small cross-section of how AI is changing the GCS centers and how we are influencing the change and growth of AI technology.



Welcome!

Welcome to the latest issue of InSiDE, the biannual Gauss Centre for Supercomputing magazine showcasing innovative supercomputing developments in Germany. In this issue, we highlight significant investments being made at our centers to advance state-of-the-art computing. With the rapid rise of artificial intelligence (AI), our centers have broadened the scope of our hardware and training offerings, as well as our allocation programs to welcome new research communities seeking to benefit from the promises offered by AI.

With the installation of JUPITER at JSC, Europe not only crosses the exaflop threshold for computing performance, the GCS research community has one of the world's most powerful machines for AI research at its disposal (page 5). While their latest machines are not exascale systems, both HLRS and LRZ have new systems that are offering users larger volumes of GPU accelerators to ensure that the AI community is well-served by the centers' respective flagship systems (page 7). All the while, the centers' staffs are working closely and sharing best practices for how to be good environmental stewards of their resources. Learn more about how GCS centers are approaching energy efficiency in the AI era on page 9.

Together with our European partners, we are focused on supporting researchers from academia and industry to ensure that this game-changing technology is developed efficiently, responsibly, and in a way that ensures Europe's competitiveness between other major technological players. In the last year, the EuroHPC Joint Undertaking has funded a series of AI Factories aimed at providing end-to-end support and world-class computational resources to European Researchers. In late 2024, HLRS was selected to lead one of the first funded AI Factories. The HammerHAI project, as it is called, also includes significant support from LRZ (page 26). Early in 2025, EuroHPC announced a second round of funding for AI factories, selecting the JAIF consortium led by JSC (page 25). These factories validate our centers' leading roles in developing AI in Europe in a sustainable, responsible manner.

Our centers have helped foster AI research from its earliest stages, and as the technology reshapes countless global industries over the next decade, we are well-positioned to balance the drive for innovation with the need for responsible stewardship of these powerful resources. This issue highlights just some of the ways that our centers are actively sculpting the future roles of AI in our lives.

Prof. Dieter Kranzlmüller
Prof. Thomas Lippert
Prof. Michael Resch

Prof. Dr. D. Kranzlmüller
Director, Leibniz Supercomputing Centre
Chairman of the GCS Board of Directors

Prof. Dr. Dr. Th. Lippert
Director, Jülich Supercomputing Centre
Co-Chair of the GCS Board of Directors

Prof. Dr.-Ing. Dr. h.c. Dr. h.c. Hon.-Prof. M. M. Resch
Director, High-Performance Computing Center Stuttgart
Co-Chair of the GCS Board of Directors

Contents

JSC is building a Modular Data Centre to house JUPITER, Europe's first exascale system.

JUPITER at the Gates

JSC is preparing to launch Europe's first exascale supercomputer, offering researchers new computational power to tackle some of the most difficult research challenges of our time.

Europe's first exascale supercomputer, JUPITER, is currently coming to life at Forschungszentrum Jülich, as part of a massive logistical endeavor. The system, which will be operated by the Jülich Supercomputing Centre (JSC) as part of the Gauss Center for Supercomputing (GCS), is being installed in a Modular HPC Data Centre (MDC). The two installations run in parallel, hand in hand, where the German-French supercomputer consortium ParTec-Eviden is responsible for finishing in JUPITER, and Eviden to finish the MDC.

JUPITER will be the first supercomputer in Europe surpassing one ExaFLOP/s per second of performance, equivalent to one quintillion (a "1" followed by 18 zeros) calculations per second at 64-bit double precision and over 40 ExaFLOP/s in 8-bit precision. With around 50 containers, the MDC will be the first public, container-based data center at this massive scale. As such, it will enable seamless upgrades and adjustments of the system and its infrastructure, while keeping energy efficiency at the system's core. JUPITER will be a key driver of progress for Germany and Europe, contributing to advancing Europe's technological and digital sovereignty. The supercomputer will be able to handle classical high-performance computing (HPC) and artificial intelligence (AI) workloads at the highest scale.

JUPITER (short for "Joint Undertaking Pioneer for Innovative and Transformative Exascale Research") is funded half by the European Union through the EuroHPC Joint Undertaking (EuroHPC JU) and a quarter each by the German Federal Ministry of Research, Technology, and Space (BMBF) and the Ministry of Culture and Science of the State of North Rhine-Westphalia (MKW NRW) via GCS.

An Eventful Journey

In the past, building data centers and installing supercomputing systems of the highest performance usually took multiple years, often spanning half a decade or more. With the rapid deployment process of JUPITER and its data center, it took only about 1.5 years from contract signing to initial availability.

Already six months after the contract signing, JUPITER's first module, the JUPITER Exascale Development Instrument (JEDI), was installed in April, 2024. It features the same equipment that the final JUPITER Booster Module will have, allowing scientists to access the hardware and optimize their codes as part of the "JUPITER Research and Early Access Program" (JUREAP). This access was supervised by experts from JSC and allows everyone involved to get ready for the final system.

During the International Supercomputing Conference (ISC24) in June, 2024, JEDI was ranked first in the Green500 list of the most energy-efficient supercomputers worldwide. After accomplishing the first major milestones with the installation of JEDI, the deployment of JUPITER has been continuously progressing, as a logistical opus magnum for all involved parties.

Last September, the initial 10 logistics containers for the MDC were placed onto the base plate. Using a heavy-duty crane, the specialized modules were carefully positioned. The MDC consists of container modules using an area of 2,300 square meters and offers a flexible and cost-effective alternative to conventional solid construction thanks to significantly shorter planning and construction times as well as reduced operating costs.

Some Impressive Figures

Through the winter of 2024 and into the first half of 2025, further MDC components were delivered, including 15 transformer stations, each capable of supplying 2.5 megawatts of power. Four containers form the data hall that houses the 29-petabyte ExaFLASH, 308-petabyte ExaSTORE, and the JUPITER Service Island, including the 3.2 terabit-per-second network connection to the outside world. The core computing power is distributed across seven double containers, called IT Modules, each containing up to 20 Eviden BullSequana XH3000 racks, with 125 racks for JUPITER Booster and up to 15 racks for JUPITER Cluster. An eighth IT Module is planned for future use for additional initiatives, such as for the JUPITER AI Factory system, JARVIS, which is used for cloud-accessible inference tasks. Both the data hall and the IT Modules have corresponding cooling units on the roof of the containers.

Lightning-fast Components for HPC and AI

Building on the dynamic Modular System Architecture (dMSA), developed by JSC and other European partners through the DEEP research projects, the JUPITER system will be made up of two main compute modules, a Booster and a Cluster, supported by multiple storage modules.

The completed JUPITER Booster system with its 125 racks and 6,000 nodes will feature around 24,000 NVIDIA GH200 Grace Hopper superchips, designed specifically for high-performance simulations and AI model training. This configuration will allow JUPITER to reach over 40 ExaFLOP/s in lower-precision dense 8-bit calculations, making it one of the fastest AI systems in the world publicly available to academia and industry. This performance is achieved within a power budget of around 17 megawatts.

The Grace-Hopper superchips of JUPITER Booster are a novel, tight combination between NVIDIA's first ARM-based Grace CPU and their recent Hopper GPU. CPU and GPU can both access each other's memory coherently, allowing the GPU to expand its memory address space, such as for large-scale simulations with a high memory demand as in JUQCS (short for "Jülich Universal Quantum Computer Simulator") or AI workloads.

The general-purpose Cluster Module is designed for workloads that do not require accelerator-based computing but a high memory bandwidth instead. The Cluster will be the first system using the Rhea 1 processor, which is commercialized by the European startup SiPearl and is a result of the European Processor Initiative. The installation of the Cluster will happen a few months after the Booster Module.

Beside the successful JUREAP program, with more than 120 applications and over 30 national and international lighthouse applications, the very first call for applications from research groups and companies to use JUPITER was issued from January to March 2025. In addition, a dedicated "AI Competition" is being offered in the first months of system operation. As quickly as the supercomputer is being completed, the desire to use it is growing.
Maximilian Tandi, Benedikt von St. Vieth

For more information about the Gauss AI Compute Competition, please visit: https://www.fz-juelich.de/en/ias/jsc/news/news-items/news-flashes/gauss-ai-compute-competition-en

More about JUPITER: https://www.fz-juelich.de/en/ias/jsc/jupiter

More about JUPITER tech: https://www.fz-juelich.de/en/ias/jsc/jupiter/tech

Delivery of an IT Module built of two containers, each container housing 10 racks.

Inside view of an IT Module with 10 BullSequana XH3000 racks on each side.

Hardware Roadmaps at HLRS and LRZ Embrace Accelerated Architectures

HLRS and LRZ are bolstering access to GPU accelerators on their systems to enable more advanced artificial intelligence research.

In the first half of 2025, both the High-Performance Computing Center Stuttgart (HLRS) and the Leibniz Supercomputing Centre (LRZ) began welcoming the first users to their newest high-performance computing (HPC) systems – Hunter at HLRS and SuperMUC-NG Phase 2 at LRZ. Both systems are the first flagship systems at their respective centers to use graphics processing units (GPUs) to accelerate performance. Over the last decade, the HPC community has increasingly embraced GPUs to both accelerate traditional modeling and simulation and increase the computing power available for researchers working in the rapidly growing field of artificial intelligence (AI).

"The rapid development of AI and an increasing focus on sustainability in supercomputing mean that high-performance computing is currently going through an exciting, transformative period," said Prof. Michael Resch, Director of HLRS. "With Hunter, our user community gains a state-of-the-art infrastructure of HLRS, in conjunction with Hunter's inauguration in January, 2025."

"At the core of all LRZ activities is the user," said Prof. Dieter Kranzlmüller, Director of LRZ. "It is our utmost priority to provide researchers with the resources and services they need to excel in their scientific domains. Over the last years, we've observed our users accessing our systems not only for classical modeling and simulation, but increasingly for data analysis with artificial intelligence methods."

The third member of the Gauss Centre for Supercomputing (GCS), the Jülich Supercomputing Centre (JSC), is set to launch the GPU-accelerated JUPITER, Europe's first exascale system, in summer 2025, giving all three GCS centers new, powerful tools to embrace these larger hardware shifts in the HPC community. (For more information on JUPITER, please visit page 5).

While each center's path is different, all three GCS centers are focused on bolstering AI capabilities while also providing powerful, reliable systems to accelerate scientific discovery.

HLRS Inaugurates Hunter on the path to Herder

For its newest system, Hunter, HLRS partnered with Hewlett Packard Enterprise (HPE) to implement a vision that includes installing two systems over the next decade. Hunter is conceived as a transitional system to prepare for the more powerful Herder, planned for 2027. Hunter offers researchers slightly more than 48 petaflops, or 48 quadrillion calculations per second. The system nearly doubles the peak performance of HLRS's prior flagship system, Hawk, but does so using only 188 nodes. Unlike Hawk, which only used CPUs, each of Hunter's 188 nodes is based on AMD's Instinct MI300A accelerated processing unit (APU). The Instinct APU combines CPUs, GPUs, and high-bandwidth memory into a single node, meaning that this acceleration also comes with a smaller spatial and energy footprint.

Perhaps most importantly, Hunter offers HLRS users the ability to optimize their application codes to run efficiently on a GPU-based architecture. This will be essential for the jump to Herder, which will offer a peak performance of several hundred petaflops, largely through its use of GPU accelerators. This means that long-time HLRS users will have three years to work closely with the center's user support staff, attend training courses to expand their knowledge of GPU programming, and generally refine their workflows to take advantage of the leap in computational performance.

Because GPUs are also essential for large-scale data analysis and AI, Hunter holds the potential to open up resources to new user communities. The center has already partnered with the Stuttgart-based AI consulting firm Seedbox.ai, for example, which began using Hunter even before it officially entered production to train new large language models in 24 languages.

SuperMUC-NG Phase 2 accelerates applications in advance of Blue Lion

While HLRS partnered with HPE to build its next two systems, LRZ has worked with Intel and Lenovo to build and then expand its SuperMUC-NG system over the last seven years. This year, the center is finishing acceptance testing for the system's latest expansion, complementing SuperMUC-NG with its Phase 2. Once accepted, the system will offer roughly 28 petaflops and will be accelerated with 960 Intel Ponte Vecchio GPUs.

Like Hunter, SuperMUC-NG Phase 2 offers LRZ users GPUs to accelerate simulation work while also providing a powerful boost for AI workloads. The center already invested in GPU-enabled envionments with LRZ's AI systems and Bayern KI, and increased its user support staff to accommodate more diverse research using AI methods, but SuperMUC-NG Phase 2 provides the largest GPU-accelerated system to date.

In late 2024, LRZ announced a partnership with HPE for its next-generation system, Blue Lion. Scheduled to be deployed in 2027, the flagship system will offer researchers a massive performance increase – the center anticipates that it will be at least 30 times as powerful as SuperMUC-NG Phase 2 – by greatly expanding GPU processing power in the system. Blue Lion will use next-generation NVIDIA GPUs and HPE's "Slingshot" interconnects to enable training large models for AI applications.

"It takes a lot of work to go through the full procurement process of a new supercomputer, but it is also immensely exciting," said Kranzlmüller. "It allows us to look into the future of supercomputing. The anticipation is growing about how the scientific community will use this system to accelerate their progress toward new discoveries."

> "GCS is one of Europe's HPC leaders, and as AI technologies rapidly evolve, our centers are well-positioned to support this new era of computing technologies."
> Dr. Claus-Axel Müller, GCS

Building AI ecosystems through collaboration

As HLRS and LRZ scale up the hardware available for AI research, the centers' staffs are also building out an AI-optimized system as well as software solutions and support needed to reach new frontiers in AI research. In late 2024, the EuroHPC Joint Undertaking announced a first round of funding for "AI Factories," launching investments in AI-optimized systems and expanded resources and support for artificial intelligence across Europe. The HammerHAI consortium – led by HLRS and including LRZ – was among the first seven AI Factories selected. (For more information on HammerHAI, please visit page 26). With a project budget of €85 million, HLRS will host a new AI-optimized supercomputer and the consortium will grow its support staff to advance Europe's AI uptake – by growing European AI expertise, providing infrastructure to enable the development of "made-in-Europe" AI technologies, and lowering the barrier of access to large numbers of GPUs for industry and academic researchers.

Dr. Claus-Axel Müller, Managing Director of GCS, pointed out that these investments in hardware, software development, training, and support align closely with GCS's mission to both support AI innovation and maintain world-class computational resources for simulation. "GCS is one of Europe's HPC leaders, and as AI technologies rapidly evolve, our centers are well-positioned to support this new era of computing technologies while delivering on our core mission of enabling essential scientific research for our diverse user community," he said. *Eric Gedenk*

GCS Centers Embrace the AI Era While Focusing on Energy Efficiency

Artificial intelligence is increasingly integrating into workflows at both public and private computing centers. As its centers scale up their AI offerings, GCS is putting sustainability and responsible stewardship of their systems at the fore of its long-term investment and infrastructure plans.

Artificial intelligence (AI) is revolutionizing how we work, study, build, and plan. While this computing paradigm shift is opening entirely new areas of research and promises rapid innovation across many industries, its reliance on power-hungry GPUs and need for long runtimes to train data are driving up energy usage at high-performance computing (HPC) centers.

High-performance computing (HPC) centers are among the earliest laboratories to foster AI development and growth. As this revolutionary, disruptive technology rapidly grows and integrates into new corners of our lives, the three Gauss Centre for Supercomputing (GCS) centers are embracing changes to their systems and processes to better accommodate research on AI systems.

These changes include how the centers' staffs approach system procurement, training, system access, and user support, among other core functions. As the computer rooms at the High-Performance Computing Center Stuttgart (HLRS), Jülich Supercomputing Centre (JSC), and Leibniz Supercomputing Centre (LRZ) fill with new, diverse systems, the centers are ensuring that they maintain a focus on sustainability.

"There are multiple levels we have to address with regard to energy efficiency," said Benedikt von St. Vieth, Head of JSC's HPC, Cloud and Data Systems and Services division. "One is just the bare infrastructure: we try to reduce energy and water consumed in operating the system to lower the environmental impact. Second, we focus on hardware, where we see a significant change from classical

LRZ has long used warm-water cooling to improve the energy efficiency of operating its supercomputing systems. Its latest system, SuperMUC-NG Phase 2, the power supply units are also directly cooled with warm water, allowing the center to nearly create room-neutral system operation.

CPU-based systems to accelerator-based systems. While energy demands for large systems are still high, we are improving energy-to-solution. Finally, we have software, developing better scheduling systems and tools that can help users to monitor and reduce their energy footprint when using these systems."

Infrastructure innovations focus on energy savings, resource reuse

While traditional HPC systems are indispensable research tools, they have always required a lot of energy to operate. For over a decade, GCS centers have built out infrastructure teams that actively seek new ways to reduce energy costs and identify opportunities to reuse waste heat generated by large supercomputers. As the centers complement their flagship systems with new clusters optimized for AI workflows, staffs are developing new ways to further reduce computing centers' environmental impacts.

One of the basic infrastructure changes starts long before any new computing racks are installed: procuring energy-efficient systems. All three GCS centers have gone through open procurement processes for new, flagship systems within the last three years.

"For us at LRZ, sustainability is an exclusion criterion in our system procurement process – so it is a "must" when submitting a proposal," said Dr. Herbert Huber, Head of LRZ's High-Performance Systems division. JSC and HLRS have also increased the emphasis on sustainability in their procurements in recent years and use energy efficiency as one of the criteria in evaluating which technology provider to partner with for new systems.

Designing an energy-efficient system goes beyond just buying the right system components, though – it also means focusing on how to improve the efficiency of the infrastructure supporting the system. Huber has helped spearhead LRZ's early embrace of liquid cooling technologies for HPC systems, pivoting the center away from using fans and chiller systems over the last 25 years.

In the SuperMUC-NG Phase 2 supercomputer, currently going through system acceptance, the computing nodes' power supply units are directly cooled with warm water, which means that the compute node racks can be thermally insulated. This allows the center to achieve 98 percent efficiency of its direct warm-water cooling, essentially creating virtually room-neutral operation of the system. JSC's JUWELS system has used direct warm water cooling since 2018, and the latest system, JUPITER – set to be Europe's first exascale system when its installation concludes in the first half of 2025 – as well as HLRS's newest system, Hunter, also use direct warm-water cooling.

Direct warm-water cooling also offers another bonus: the ability to reuse waste heat. Over the last five years, all three centers have started to work with their home campuses to build waste heat loops that can heat offices and buildings near their computing centers. LRZ began heating its office and lecture buildings with the waste heat of its supercomputers in 2012. In addition, LRZ is in discussion with the infrastructure team of Technical University of Munich (TUM) about supplying heat for the Garching campus that houses LRZ and other research institutions. Forschungszentrum Jülich (FZJ) is also planning to use JUPITER's waste heat in buildings across the FZJ campus and beyond. In advance of its next system, Herder, HLRS is constructing a new building and working with the University of Stuttgart to incorporate waste heat use in nearby buildings on campus.

Hardware advancements improve power-to-solution

Even before the AI boom, HPC centers began embracing hardware changes that improved energy efficiency. Since the HPC community started incorporating GPU acceleration into systems in 2012, researchers have needed to modify applications and community codes to take advantage of this new technology for parallel computing. As more research communities have embraced GPUs, HPC centers have, too. While GPUs generally have a higher power density than CPUs – meaning they use more power at max performance – designing hybrid CPU-GPU systems allows researchers to offload repetitive, time-intensive calculations onto GPUs. This division of labor helps to reduce time-to-solution in many applications, saving energy in the process.

While JSC experimented with smaller numbers of GPUs in previous systems, the center built JUWELS in 2018 as a modular supercomputer that would incorporate a GPU-centric module – the JUWELS Booster. This strategy made it possible to accelerate traditional modeling and simulation applications while simultaneously offering researchers interested in machine learning and other early AI applications a system architecture more suited to their needs. JSC's new JUPITER system makes use of 24,000 of NVIDIA's Grace-Hopper superchips (GH200), which offers tight integration between NVIDIA's Grace CPU and its Hopper GPU.

Shortly after JUWELS' installation, LRZ and HLRS also began offering their users access to GPU clusters. In SuperMUC-NG Phase 2 and Hunter, both installed this year, the centers ramped up the number of GPUs available on their flagship systems. When it completes acceptance testing later in 2025, SuperMUC-NG Phase 2 will offer users 960 Intel Ponte Vecchio GPUs. Through its relationship with Hewlett-Packard Enterprise (HPE), HLRS is using AMD chips in its newest system, Hunter, as well as its future system, Herder. Hunter uses 188 AMD Instinct MI300A accelerated processing units (APU), which combine CPUs, GPUs, and high-bandwidth memory in a single chip.

"With Hunter's architecture, we are getting almost double the performance of our previous Hawk supercomputer but use 80 percent less energy when we compare the systems' peak performances." Dr. Ralf Schneider, HLRS

"With Hunter's architecture, we are getting almost double the performance of our previous Hawk supercomputer but use 80 percent less energy when we compare the systems' peak performances," said Dr. Ralf Schneider, Head of HLRS's Numerical Methods and Libraries team.

Software tools empower users and staff to further improve energy efficiency

No matter how much planning for sustainability goes into infrastructure and hardware choices, how a computing system is used during day-to-day operation has a significant impact on its power usage. As energy costs in Germany have risen and demand for powerful GPUs grows for AI applications, GCS centers' staffs have developed new software capabilities to ensure these resources are used as efficiently as possible.

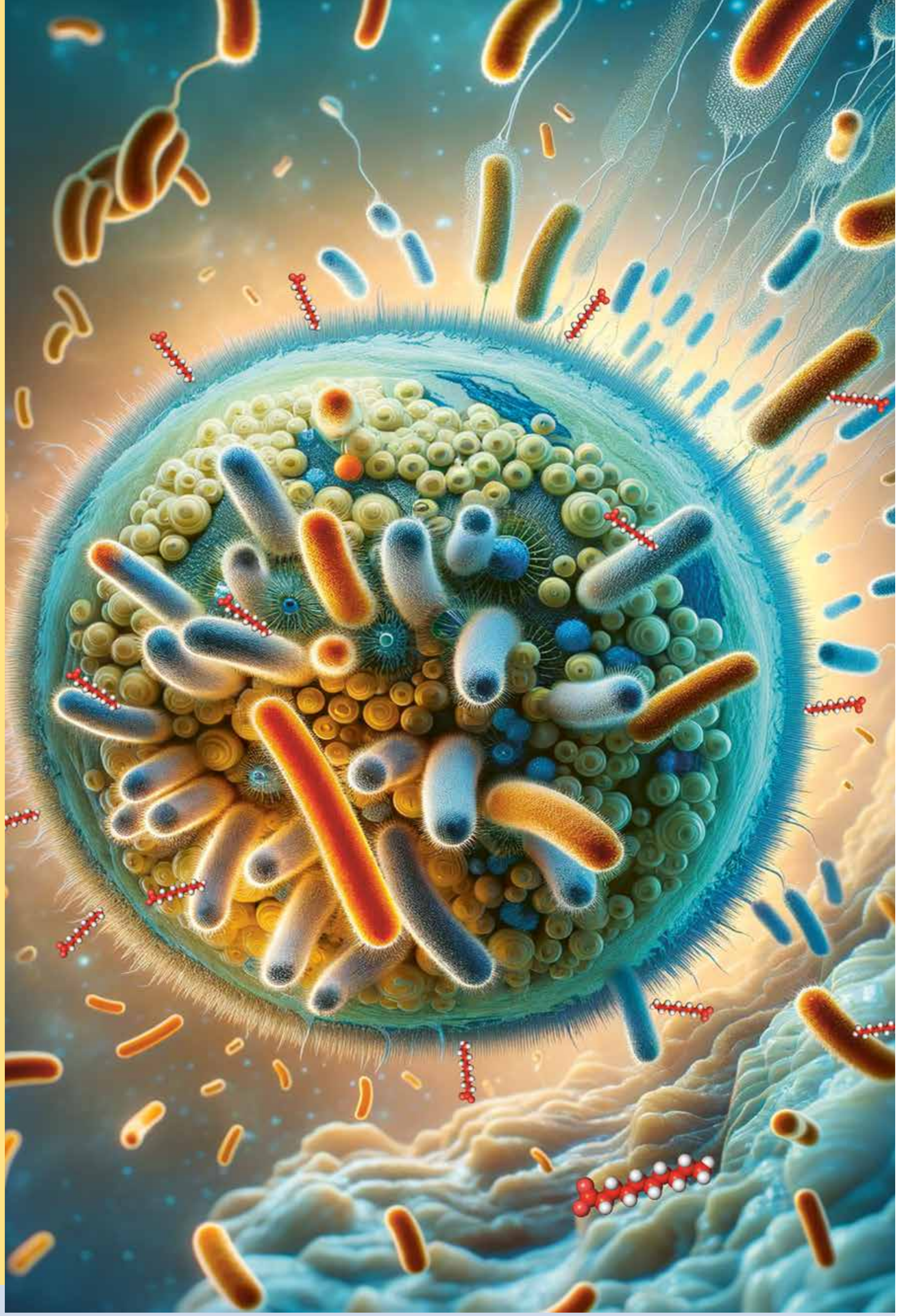
HLRS has been working with Hewlett-Packard Enterprise – manufacturer of the Hawk, Hunter, and Herder systems – since 2020 to improve how the center monitors and improves applications' power usage. Together, they developed a dynamic power management solution that was implemented in 2024 on Hawk that closely monitors applications' performance and dynamically caps how much energy gets used on them.

HPC centers have already experimented with so-called "static" power capping, which limits the maximum power used by a supercomputer. While this approach can save energy during times when a machine is sparsely used, it also lowers processors' performance, extending the time-to-solution for applications running on the system.

The dynamic power management tool identifies "memory-bound" applications – applications whose performance is limited by available memory and data transfer speed across nodes rather than raw processing power – and throttles down processors that would otherwise be idling. Schneider and his collaborators ran a comparison in December 2024 to compare power consumption when Hawk was not power capped and found that dynamic power management reduced energy use by 20 percent without impacting application performance. "It may not seem like a large savings when we are talking about saving a certain percentage of power used per node, but it winds up being a large amount of electricity," Schneider said. "This is taxpayers' money and CO_2 emissions we are saving."

The other GCS centers are also using power capping approaches. Huber noted that LRZ uses EAR – the Energy Aware Runtime software – to automatically clock down the processor frequencies of nodes running application program routines that do not benefit from high processor frequencies. EAR reduces the annual power consumption of SuperMUC-NG without any impact on application performance by about 30 %, and the center will extend that effort to its future AI-focused systems. At JSC, von St. Vieth also discussed how efficiently scheduling jobs and power capping would play a big role in reducing JUPITER's power consumption. "On JUPITER's test platforms, we've already been experimenting with limiting power consumption for certain applications, and we want to make sure our users are aware of these functionalities so they are aware of how their application performs."

In a time of transition in the HPC community, embracing the potential of AI also means being a responsible steward of large computing resources. Designing infrastructure that reduces and reuses waste, operating energy-efficient hardware, and developing sustainability-focused software solutions represent the GCS's approach to remaining a green-IT leader in the AI era.
Eric Gedenk



The membrane-bound phospholipase A1 from *Pseudomonas aeruginosa* is a potential drug target. PlaF remodels membrane glycerophospholipids, influencing virulence-associated signaling. Medium-chain free fatty acids, products of PlaF action, inhibit its activity. Figure generated by Rocco Gentile from a hand drawing that was refined with GPT-4.

Computational Scientists Open New Front on Antibiotic-Resistant Bacteria

Researchers at Heinrich Heine University and Forschungszentrum Jülich are using high-performance computing in tandem with experimental research to better understand a common, antibiotic-resistant infection prevalent in healthcare settings.

While drug development has helped humans overcome countless illnesses over the decades, certain viruses and bacteria have remained stubborn adversaries. In hospitals and other healthcare facilities, doctors and nurses must vigorously protect immunocompromised patients from several antibiotic-resistant infections. One of them, *Pseudomonas aeruginosa*, is a bacteria well-suited for spreading so-called nosocomial infections, or infections that most readily spread in healthcare environments among patients with weakened immune systems.

"*P. aeruginosa* is one of the candidates that the World Health Organization considers among the more threatening bacteria," said Prof. Holger Gohlke, who heads Heinrich Heine University's (HHU's) Computational Pharmaceutical Chemistry and Molecular Bioinformatics group. "Our experimental collaborators research this bacterium in a Level-2 safety lab. As a healthy person, you can enter the lab with little risk, but for immunocompromised people, it can cause severe infections."

Gohlke and his experimental collaborator, Prof. Karl-Erich Jäger at HHU's Institute of Molecular Enzyme Technology at Forschungszentrum Jülich, were funded by a German Research Foundation Collaborative Research Centre grant to focus on studying membrane dynamics in cellular systems. Their goal is to improve the drug development pipeline and better combat challenging infections like those brought about by *P. aeruginosa*. Gohlke and his research group use high-performance computing (HPC) resources at the Jülich Supercomputing Centre (JSC) to run molecular dynamics simulations aimed at better understanding one of the bacterium's enzymes responsible for cell membrane damage. This research aims at improving our understanding of *P. aeruginosa* and, in the medium term, discovering new methods to bolster our pharmaceutical defenses against antibiotic-resistant bacteria. The team's research made the cover of *JACS Au*.

Fatty acids provide clues to combating antibiotic resistance

One of the main characteristics of *P. aeruginosa* that makes it dangerous to human health lies in one class of enzymes. Specifically, so-called type-A phospholipases, or PLA1, are enzymes in *P. aeruginosa* that damage healthy cell membranes and disrupt the signaling networks cells use to combat infection. Gohlke and his collaborators identified a specific enzyme, PlaF, as one that can make the bacteria more dangerous for immunocompromised patients.

In prior research, the team had identified how certain types of fatty acids – specifically, medium-chain free fatty acids (FFAs) – could blunt PlaF activity in cells, but did not fully understand the specific reasons for its success in inactivating PlaF.

To better understand how FFAs can combat PlaF's role in increasing *P. aeruginosa's* virulence, Gohlke, Jäger, and their team used a combination of highly complex molecular dynamics simulations in concert with *in vitro* and *in vivo* experiments. With access to the JUWELS supercomputer at JSC, the researchers uncovered that FFAs not only indirectly influenced PlaFs by alternating its shape, but also by directly impacting the enzyme's center, which might lessen its influence on nearby healthy cells. In both cases, the researchers opened a potential path for molecular

Certain antibiotic-resistant bacteria such as *Pseudomonas aeruginosa* are not a large risk for healthy people, but the immunocompromised and those in hospitals can develop life-threatening "nosocomial" infections.

biologists, chemists, and other scientists to develop new pharmaceuticals to combat *P. aeruginosa* infections.

"The research community has known that fatty acids can have functions as signaling molecules," Gohlke said. "This is one strong example of the role that they can play in regulating protein activity by influencing its configuration and very directly by serving as a blocking agent for active sites in an enzyme. This is another factor to consider in the future as we look at similar bacterial systems when we are investigating new ways that pharmaceuticals can fight off infection."

HPC resources improve drug development pipeline

Gohlke has a joint appointment at JSC's parent organization, Forschungszentrum Jülich. He has also been a long-time user of JSC's computational resources. For him, having a long-running, collaborative relationship with JSC staff has provided a positive boost for his research.

"JSC has always provided our team excellent user support," Gohlke said. "Oftentimes, PhD students or postdoctoral researchers may be accessing these systems for the first time, and we feel confident in their abilities to help us with troubleshooting our application on their system whenever we need it."

Gohlke's team primarily uses the Amber code for its research, which was one of the first applications of its kind to embrace the use of GPUs in supercomputers. JSC hosted some of Europe's first GPU-accelerated systems, and as the center moves from its current flagship machine, JUWELS, to its next-generation system, JUPITER, there will be even more GPU-accelerated computing at researchers' disposal.

He indicated that the team often uses small numbers of GPUs in its simulations, but that they need to run many iterations of their simulations for meaningful insights. JUPITER will have more GPU nodes available to researchers, meaning that Gohlke's team and other researchers will have more nodes available for their research.

With its increased access to GPUs, the team is also exploring ways that machine learning methods might be able to support its drug discovery research more generally. Gohlke indicated that the team was already looking into ways to use machine learning to more effectively predict enzyme function and better judge how certain mutations could either help or hinder particular treatment methods.

Since the *JACS Au* paper came out, Gohlke's team has continued to leverage HPC resources to seek out small molecules that could ultimately impact PlaF specifically and PLA1 enzymes generally.

"We have identified compounds that show promise for positively impacting these systems," Gohlke said. "I would not claim that we have new antibiotics out of this research, but we do already have promising compounds that can act as a template for further development.
Eric Gedenk

Catching Runaway Electrons in Fusion Reactions

A team at the Max Planck Institute for Plasma Physics uses high-performance computing to address a key challenge in making fusion power generation safe and efficient.

Since the 1950s, scientists have sought to harness the power of nuclear fusion. When two atomic nuclei collide at high speeds, they fuse into a single nucleus that has less mass than the sum of two original nuclei and in return releases energy. By harnessing energy sloughed off during this rection and efficiently perpetuating it on a large enough scale, humanity could one day build fusion reactors that would offer a source of electricity that is both carbon-free and that uses a less volatile and less dangerous process than today's nuclear-fission-based power plants.

Researchers at the Max Planck Institute for Plasma Physics (IPP) are currently focused on one of the largest remaining obstacles to making fusion practical: effectively confining the ultra-hot plasma needed to sustain a large-scale fusion reaction inside donut-shaped devices called tokamaks.

"The quest for fusion energy is a massive collaborative effort and the challenges involved go well beyond problem of disruptions stemming from loss of plasma confinement," said Dr. Matthias Hölzl, Group Leader of the Non-Linear Magnetohydrodynamics (MHD) group at IPP. "That said, disruptions present one of the leading concerns for the feasibility of the tokamak concept as a whole. Understanding how to avoid and mitigate disruptions is essential if we aim to build reactors that could actually power our lives."

Hölzl and his collaborators are using high-performance computing (HPC) resources at the High-Performance Computing Center Stuttgart (HLRS) to focus on a particular phenomenon that makes plasma disruptions so problematic: highly energetic runaway electrons that may arise because of the loss of plasma confinement and cause large, localized heat loads on the walls.

Charting the paths of runaway electrons in search of better confinement

Creating a large-scale fusion reaction is no small feat: it is essentially trying to mimic the sun. Researchers confine and heat plasma material in a tokamak until it gets hot enough for fusion reactions. "For fusion to occur, the fuel needs to heat to around 100 million degrees Celsius," said Hannes Bergström, doctoral student at IPP and first author of a refereed journal article that describes the newest findings. "At these extreme temperatures, electrons are

> "Understanding how to avoid and mitigate disruptions is essential if we aim to build reactors that could actually power our lives."
> Dr. Matthias Hölzl, IPP

> "Input from our work will help make future large fusion devices more reliable and more efficient."
> **Dr. Matthias Hölzl, IPP**

no longer bound to atomic nuclei, so we say that material is in a plasma state. We use magnetic fields to confine plasma, in part by running a current through it, but the entire system is a delicate balance, and instabilities can appear that may rapidly drop the temperature and disrupt the reaction."

In these extreme environments, these liberated electrons can accelerate almost to the speed of light, becoming "relativistic" particles in the process. Via collisions, these electrons can create an avalanche with more and more of these highly energetic particles. Such a beam of high-speed particles can pose a risk to the reactor wall.

Due to the risk of damaging machine prototypes through experiment, plasma physicists rely on simulations to help understand how runaway electrons begin and how they could be averted or mitigated. To simulate plasma accurately, researchers use MHD, which treats plasma as a continuous object rather than a collection of individual particles. But the runaway electrons are not described accurately enough by such methods, and a so-called hybrid fluid-kinetic approach is needed. This approach captures the mutual interactions between the background plasma and the relativistic particles. The team was the first to develop such a model.

To get an accurate simulation of these complicated processes, researchers require HPC resources to efficiently run calculations needed to account for how the various properties during a reaction influence one another. "We have to simulate the evolution of temperature, particle transport, runaway electron generation, as well as electric and magnetic fields, and all of these things influence one another," said Bergström. "In addition, we have to model the entire tokamak realistically in 3D and account for differences in time scales – disruption to plasma can happen over seconds, but runaway electrons can start to be lost in a matter of microseconds."

The team uses its JOREK MHD code to model fusion plasma dynamics, but until recently, had been unable to use JOREK to accurately model runaway electrons from first principles, relying instead on less-accurate fluid models to include them in the simulation. As part of his PhD work, Bergström developed a model to include in JOREK that would both accurately simulate runaway electrons' behavior and couple those reactions to the general plasma dynamics in a JOREK simulation. Using the Hawk supercomputer at HLRS, the team was able to verify this additional model's accuracy, creating a simulation of a runaway electron beam hijacking the plasma current during a disruption, then converting back to being carried by the cooler plasma after the loss of the runaway electrons to the walls.

"With this new model, we can use JOREK to study the interaction between runaway electrons and the rest of the plasma with accuracy that previously just wasn't possible," Hölzl said. "We benchmarked it to make sure the model behaves as we would expect, but now we need to simulate more realistic scenarios with larger plasma volumes and longer time scales, where earlier models might not be able to tell us what we should expect to see."

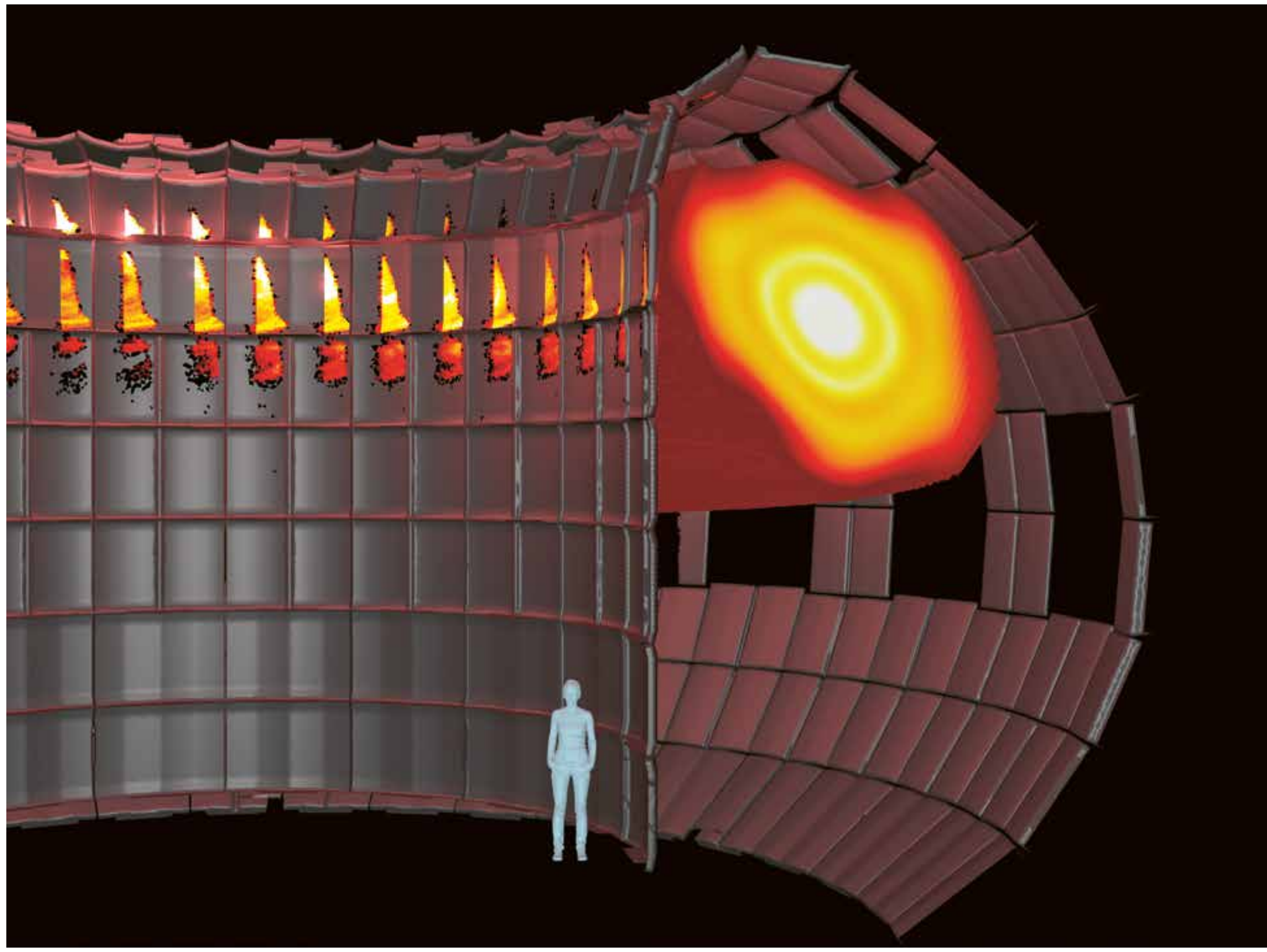
Future fusion research flows through next-generation HPC

With this new model integrated into JOREK, the researchers have set their sights on further developing the code to take advantage of the newest generation of HPC systems at HLRS and the other GCS centers – namely, they are porting their application to run on GPUs and AMD's accelerated processing units installed in HLRS's latest system, Hunter. "We expect this work to yield a substantial speed up and allow us to cross longer time scales – this is crucial if we want to study aspects of the runaway electron process that are happening across widely different timescales," Hölzl said.

With access to these accelerated systems, Hölzl hopes to continue pushing the frontier of fusion simulations from emulating conditions in a tokamak built for fundamental fusion research to a device like ITER – a large experimental fusion reactor being built in France that aims to create the largest-ever sustained fusion reaction when it goes into operation in the early 2030s. Hölzl noted that while accurately simulating a reactor on the scale of ITER with the full complexity of the new model is still beyond current capabilities, the team's work is laying the groundwork for next-generation simulations to provide insight into next-generation tokamaks.

"Input from our work will help make future large fusion devices more reliable and more efficient," he said. "The size of future devices does increase the computational cost of these simulations by orders of magnitude. That means we must focus on model development, validation, and optimization. The availability of powerful computational resources remains essential." *Eric Gedenk*

Related publication
Bergström et al. (2025) *Plasma Phys. Control.* Fusion 67 035004.
DOI: 10.1088/1361-6587/adaee7



Snapshot taken during the termination of a beam of runaway electrons in the upcoming ITER tokamak (with person for scale), caused by the growth of a plasma instability. During this violent event, the shape of the beam becomes distorted, as seen on the right. Eventually the runaways become deconfined and the resulting heat deposition on the first wall, seen on the left, is heavily localized.

SuperMUC-NG Supports Scientists' Pursuit of New RNA-Based Therapies

Researchers at the Ludwig-Maximillians-Universität München have spent over a decade researching new ways to effectively deliver RNA-based drugs to the correct place in the body. The team recently used LRZ supercomputing resources to improve its understanding of the role nanoparticles play in making sure these therapies make it to their destination.

The field of RNA-based medicine has grown since the 1990s, with researchers making strides in fighting cancer using RNA for over a decade. The technology came into popular consciousness during the COVID-19 pandemic, where the first vaccines created to combat the virus used messenger RNA (mRNA). As researchers continue to better understand how to use RNA to fight disease, they've explored how to use it to treat new conditions as well as how to improve its ability to arrive where it is intended in the body.

"You cannot just put RNA in your body through a capsule or a vaccine, because it will not end up in your cells – you need to design ways to get it there," said Dr. Benjamin Winkeljann, a group leader within the Drug Delivery Chair led by Prof. Olivia Merkel at Ludwig-Maximillians-Universität München (LMU). "We use nanocarriers that can essentially bring the RNA into a cell. More of the therapeutic entering the cell makes a medicine more effective and getting it into the correct, diseased cells specifically lowers the risk or severity of side effects for patients."

Winkeljann has spearheaded efforts to incorporate high-performance computing (HPC) simulations into the Chair's larger efforts to improve drug delivery in the body. Recently, Winkeljann and his collaborators used SuperMUC-NG at the Leibniz Supercomputing Centre (LRZ) to run molecular dynamics simulations capable of guiding the team's efforts in better understanding how certain classes of nanoparticles can help safely guide RNA therapies to their target. Working together with LRZ, the team recently developed computational methods to more accurately simulate these interactions between RNA and nanoparticles.

Molecular dynamics simulations chart the course to improved drug delivery

Unlike DNA – a double-stranded molecule that imparts important genetic information to human cells about their role and function in the body – RNA is a class of molecules that either carry out specific cell functions or create proteins to do it for them.

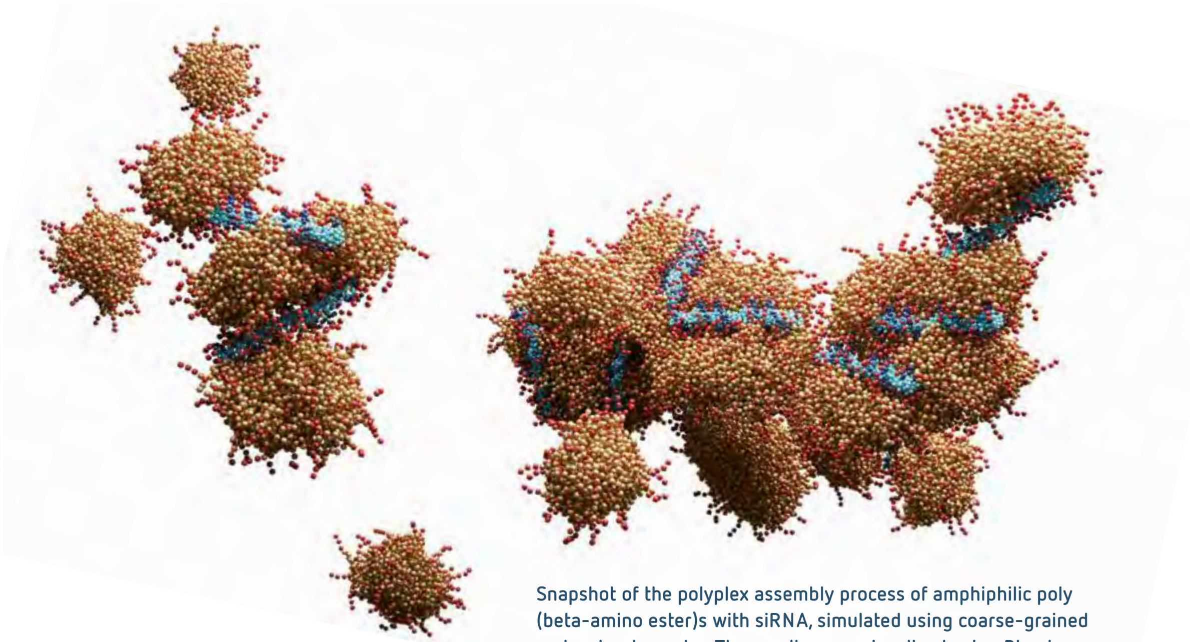
Researchers in the late 90s discovered a specific class of RNA, called short-interfering RNA (siRNA), which limits or blocks gene expression – essentially, the process by which the constituent building blocks of DNA and RNA, called nucleotides, produce proteins that give instructions to certain cells.

When it comes to genes, there can be too much of a good thing. "Certain pathological conditions are caused when genes get overexpressed, which leads to the symptoms we associate with a specific disease," Winkeljann said. "This can happen in severe asthmatic disease, COPD, or certain types of cancers." Winkeljann continued, noting that one of the primary benefits of designing siRNA therapies is its ability to attack illness at the root cause. "You can use these kinds of therapies very early in the onset of a disease rather than trying to interfere and correct for damage already done," he said.

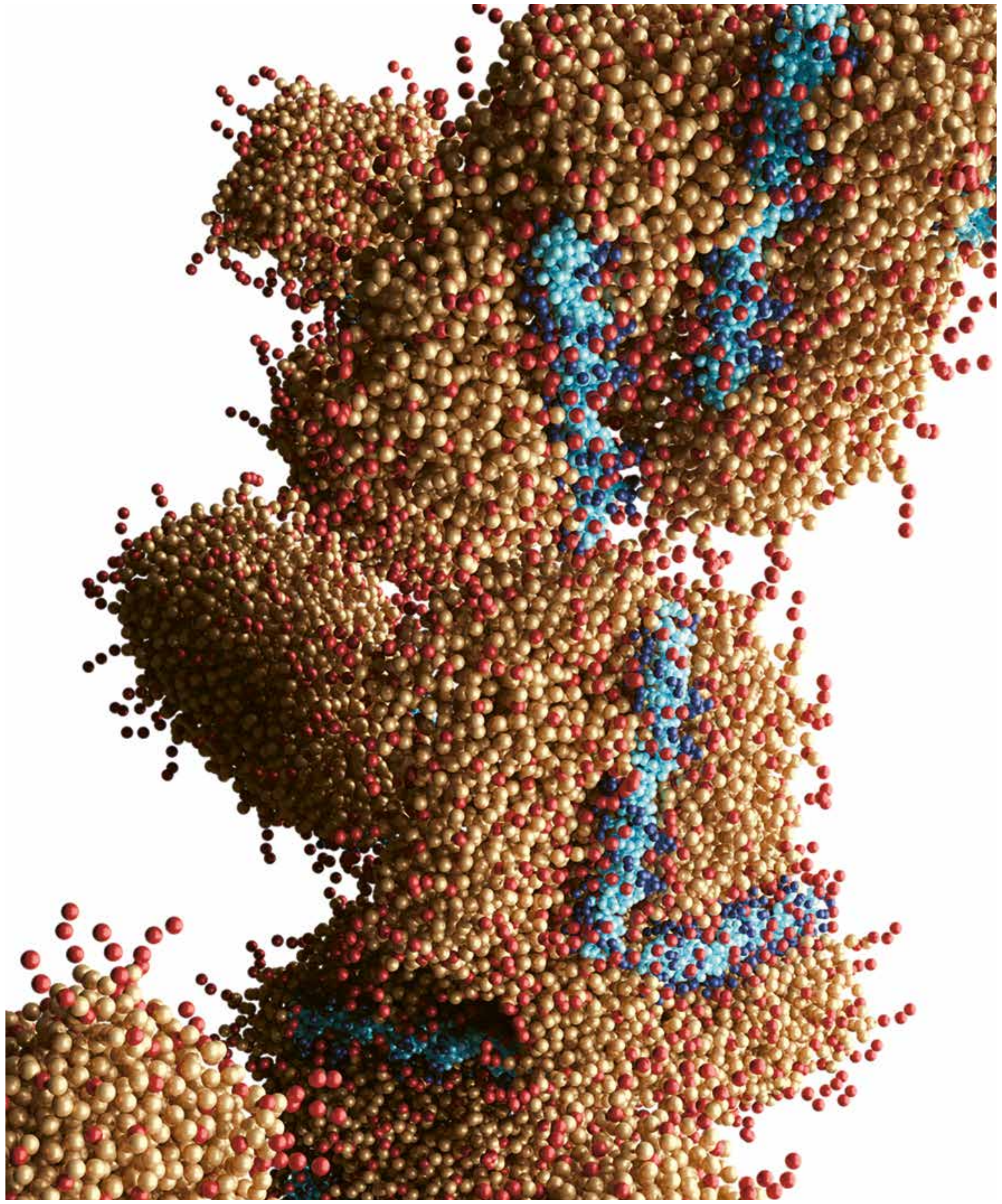
> "Certain pathological conditions are caused when genes get overexpressed, which leads to the symptoms we associate with a specific disease."
> Dr. Benjamin Winkeljann, LMU

While siRNA is well-suited to limit or silence excessive gene expression, it is not adept at finding its way to the right place to do it. RNA in medications can degrade before it ever reaches the correct place in the body, so those developing RNA-based therapies must combine the RNA with other carriers that can serve as delivery vehicles. Thus far, the several approved RNA medications on the market rely on lipids, or fat molecules, as their delivery agents, but Winkeljann and his collaborators wanted to investigate whether they could get better results using cationic polymers. Polymers are large molecules with consistent, repeating structures – DNA and RNA are both subcategories of polymers. Positively charged cationic polymers are attracted to negatively charged RNA molecules and form so-called "polyplexes" that can safely guide the vulnerable RNA to its intended location.

When Winkeljann joined the LMU Drug Delivery Chair in 2021, the team already had ample experience working on developing RNA therapies, but he sought a way to augment the iterative trial-and-error experimental process. He focused his postdoctoral work on integrating simulation into the team's workflow. "I thought it would be interesting to involve computational methods in our research, and it



Snapshot of the polyplex assembly process of amphiphilic poly (beta-amino ester)s with siRNA, simulated using coarse-grained molecular dynamics. The results were visualized using Blender and computed on the SuperMUC-NG system at LRZ.



Using coarse-grained molecular dynamics simulations on SuperMUC-NG at LRZ, Winkeljann and his collaborators have a powerful new tool for developing new siRNA-based therapies.

also felt like I was able to challenge myself with a new skillset, as I had not worked with molecular dynamics simulations before."

To study how drugs interact in the human body, researchers use molecular dynamics (MD) simulations. MD allows scientists to observe the movements of various molecules in a system and chart how that system evolves and changes over time. While this method is accurate in recreating molecular interactions, it is also computationally expensive when trying to simulate more than a few molecules interacting for a short period of time.

Because he knew the team needed to run larger simulations to accurately model polyplex formation and interactions between multiple polyplexes, Winkeljann worked closely with LRZ user support staff to develop a computational workflow using "coarse-grained" molecular dynamics (CGMD). Instead of simulating every atom from first principles, CGMD allows researchers to group certain atoms and model them as a single "interaction site." While this approach makes some assumptions about parts of the molecular system, it also allows researchers to run simulations efficiently while maintaining good accuracy. For Winkeljann and his collaborators, having the ability to run many good simulations of a larger system is more valuable than running one or two simulations with fewer particles from first principles. "In order to understand patterns and concepts present in these complex systems, we need to be able to change parameters and run suites of simulations instead of simulating one gigantic system," he said.

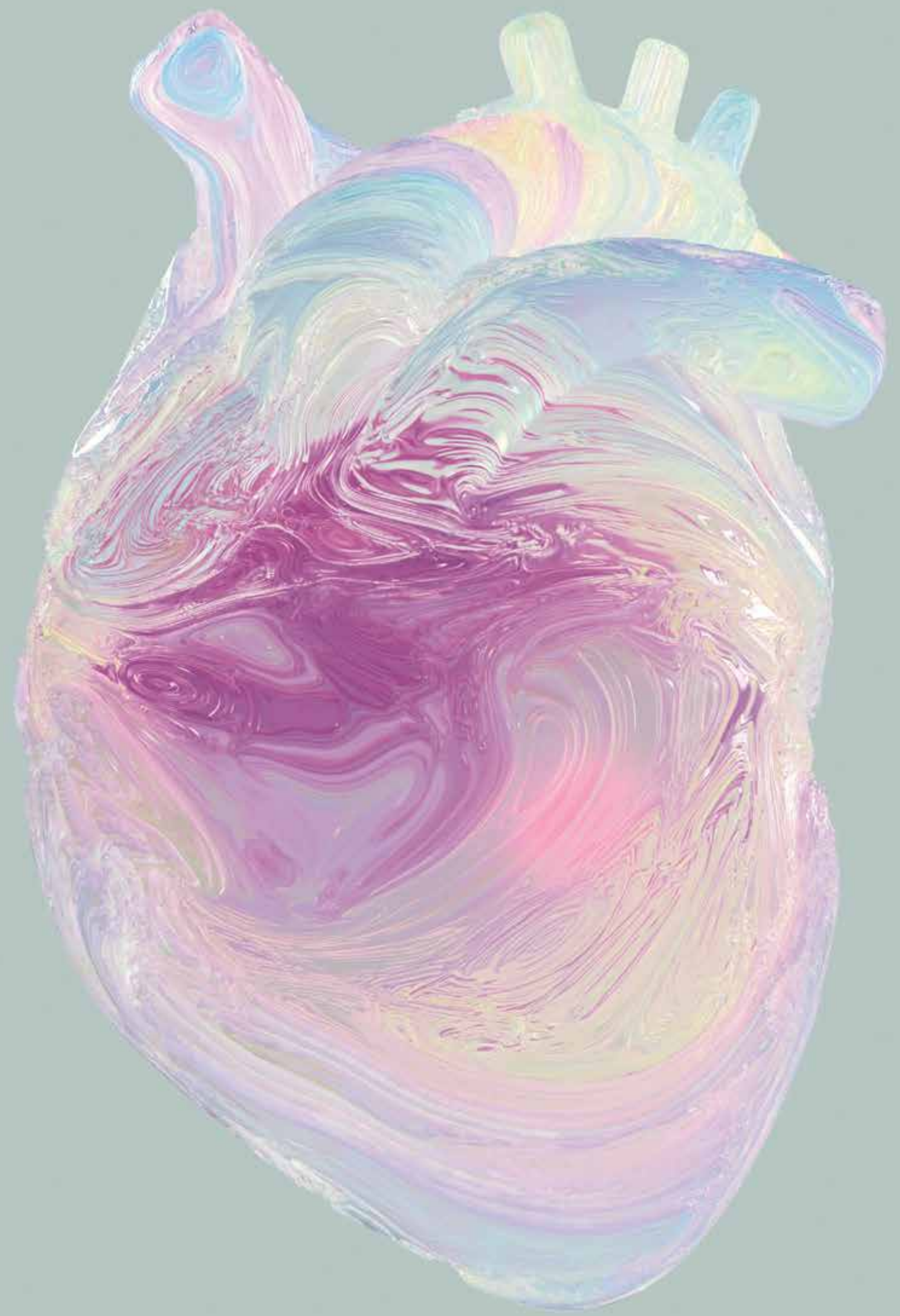
Using SuperMUC-NG, the researchers created simulations modeling the molecular organization and dynamics of siRNA polyplexes and compared those results with experimental findings. Its simulations showed good agreement with their experiments, providing a new research and validation tool for use in the development of new siRNA therapies. The team published its results in *Nano Letters*.

Public HPC centers support fundamental research underpinning the drug discovery pipeline

With a trusted CGMD method now in its toolbox, the team can complement its experimental work with insights from simulations. Winkeljann credited the close collaboration with LRZ staff in helping the team scale up its application to take full advantage of SuperMUC-NG. "I did not have extensive computational experience before joining this team, and it was a big jump to take our application from our Chair's cluster to a large system like SuperMUC-NG," he said. "LRZ helped us go through a scaling up process when we started to think about applying for time, so we had a test project on the system, then we had a regular project, and now we've got a large-scale allocation." The team was paired with Dr. Helmut Brüchle at LRZ, whose experience with MD simulations allowed the team to refine its application.

Further, Winkeljann appreciated working at a public HPC center that gives his team the freedom to do fundamental research. "Having public support versus commercially interested support allows us to maintain the independence of our research," he said. "We are doing work that is more fundamental to the field than the usual pharmaceutical formulation developer because we are digging into the underlying mechanisms of these systems. If this was privately or commercially funded research, there would undoubtedly be other goals in mind that would make it difficult to pursue this kind of research."
Eric Gedenk

Related publication
Steinegger, K. et al. (2024). *Nano Lett.*
DOI: https://doi.org/10.1021/acs.nanolett.4c04291



As part of the deal.iiX project, researchers will help bring computational tools for modelling human organs to the next generation of high-performance computing resources.

Programming Digital Twins of Organs

The program library deal.ii has a variety of tools for modelling full of mathematical tools for calculating fluid flows or modeling the deformation of geometric bodies. An interdisciplinary research group is currently elevating deal.II to the Exascale level and devel-oping programs for organ simulation.

Calculating and digitally modeling the human heart, lungs, or brain: For such tasks, the open-source software package deal.II is highly recommended. Based on the C++ programming language, it comprises around 600,000 lines of code and provides many tools for developing innovative solvers for partial differential equations. These, in turn, form the basis for calculating fluid and gas flows or modeling the deformation of solid bodies – fundamental requirements for digitally representing and simulating human organs. Martin Kronbichler, a professor of mathematics at Ruhr University Bochum, co-developed this software library and currently leads an interdisciplinary European project. Along with 12 other universities and research institutions, including the Leibniz Supercomputing Centre (LRZ), the goal is to adapt deal.II and its associated codes and applications to the nextgeneration of heterogeneous Exascale level supercomputers. "Whenever a new computer system arrives, it means a lot of work – especially adapting the code," says Kronbichler. "It's an exciting time when we, as mathematicians, can re-evaluate decisions made 10 or 15 years ago and ask whether they still lead to efficient execution on actual hard-ware – or whether we need to take entirely new directions."

Greater Energy Efficiency Through Optimized Algorithms

deal.II provides the shared mathematical tools for differential equations that allow different project teams to adapt existing programs and algorithms or build additional application-specific models. The researchers are particularly focused on tools for organ modeling and flow simulation. Based on deal.II, a new Exascale framework called "dealii-X" is being developed, specialized to create digital twins of the human body. The first project phase focuses on the algorithms that determine the sequence of computational steps when running code on parallel computer systems. "On the software side, we can continue using many features – about 80 to 90 percent of the code will still run on CPUs," explains Kronbichler. "The remaining 10 to 20 percent still leaves plenty of code and work – but it's where we can achieve the greatest speedup."
To accelerate results, heterogeneous supercomputers rely on a combination of Central Processing Units (CPUs) and Graphics Processing Units (GPUs), often with additional accelerators. Their interaction must be reorchestrated to enable more sustainable and efficient computing: "Most

> "On one hand, we're adapting code to new technologies; on the other, stronger computers now allow us to model processes that were previously impossible to calculate."
> Martin Kronbichler, Ruhr University in Bochum

> "The complexity of the underlying mathematical models has so far prevented the transfer of simulation knowledge into clinical practice."
> Martin Kronbichler, Ruhr University in Bochum

Facts & Figures: DEALII
600,000 codelines, 30 years of work
10,000 pages of documentation
45 tutorial programs
20,000 downloads/year

simulations rely on a variety of mathematical components, but they typically result in linear and nonlinear systems of equations – and those are ideal for GPUs," says Kronbichler. GPUs are optimized for fast computation and data processing, and they use less power than CPUs, which also handle control functions. "deal.II contains many algorithms that we can reformulate to rebalance computations and data access from processor cores. Usually, we reduce data access by performing redundant calculations." Data transfer accounts for a significant portion of energy consumption in supercomputing. To reduce this, the team relies on matrix-free operators in Krylov subspace and multigrid methods that replace traditional data structures.

Unlocking the Potential of New Supercomputers

In general, as computational power increases, so do researchers' expectations: They want to incorporate and model more parameters for more detailed simulations. Application specialists in the interdisciplinary dealii-X project are already optimizing programs to simulate organs and bodily functions in greater detail. dealii-X includes simulation tools for the brain, lungs, blood vessels, and liver. The project team includes not only computer scientists, mathematicians software engineers, and medical experts. "On one hand, we're adapting code to new technologies; on the other, stronger computers now allow us to model processes that were previously impossible to calculate," Kronbichler explains. Existing simulations should also be executed more quickly: "With each generation of supercomputers, it becomes harder to fully harness and leverage their technical potential." To better understand the increasingly complex development of code and algorithms, all adaptations and new developments are tested on the heterogeneous supercomputers at LRZ. These tests produce performance benchmarks that guide further adjustments, and the experience gained helps optimize high-performance computers (HPC) for specific software – by, for example, adjusting clock rates.

AI Models Complement HPC

Ultimately, dealii-X isn't only about supercomputing: some research groups are also working on methods and models for Artificial Intelligence. Statistical computations can enhance mathematical-physical models, and AI models can be trained using simulation results to quickly generate additional scenarios. "The complexity of the underlying mathematical models has so far prevented simulation knowledge from being transferred into clinical practice," Kronbichler notes. By the end of 2026, dealii-X aims to deliver tools that allow the results of mathematical-physical models – or their digital organ twins – to be integrated into software for hospitals and clinics. Digital hearts, lungs, or blood vessels could then be personalized with patient data to support therapies. *Susanne Vieser*

PROJECT	dealiiX: An Exascale Framework for Digital Twins of Humans
FUNDING AGENCY	HORIZON JU Research and Innovation Actions
FUNDING AMOUNT	€3.9 million
RUNTIME	Oct 1, 2024 to Dec 31, 2026

Europe's AI Booster: JUPITER AI Factory Brings Exascale Power to Business and Science

In the second award announcement for the EuroHPC Joint Undertaking's AI Factory initiative, JSC was selected to lead JAIF. Providing access for industry, research, and the public sector, the project is bolstering Europe's AI capabilities.

In its second round of AI Factory awards, the EuroHPC Joint Undertaking selected the JUPITER AI Factory (JAIF) consortium led by the Forschungszentrum Jülich to host one of Europe's newest AI Factories. JAIF builds on a coordinated approach in the Gauss Centre for Supercomputing in which multiple, complementary AI factories support the rapid growth of artificial intelligence in industry. The consortium consists of the Jülich Supercomputing Centre (JSC), the Center for Artificial Intelligence at RWTH Aachen University, the Fraunhofer Institutes for Applied Information Technology and for Intelligent and Information Systems, and the Hessian Center for Artificial Intelligence. JAIF also collaborates closely with the German AI service centers WestAI and hessian. AISC, the KI Bundesverband, as well as the Institute for Machine Learning and Artificial Intelligence (LAMARR) as associated partners.

The EuroHPC's AI Factory initiative is bringing together leading European computational resources, expertise, and data to promote and leverage AI technologies in healthcare, energy, climate, education, public sector and finance, manufacturing, among other key research areas. In late 2024, the EuroHPC JU announced the formation of seven initial AI Factories. Today's announcement added six new AI Factories, bringing the total to 13 across Europe.

JAIF will support European startups, small and medium enterprises, large industry, and the academic research community in integrating AI techniques into existing or emerging applications. The consortium is putting special emphasis on the EuroHPC JU's key strategic application areas and will support co-design of new AI applications. The foundational piece of technology for JAIF is JSC's upcoming JUPITER supercomputer. Set to go online in mid-2025, JUPITER will be Europe's first exascale supercomputer and one of the most powerful machines in the world for AI training applications. The consortium also plans to provide world-class individual support for all AI use cases – from user training and personal consulting to cutting-edge services in data curation.

JAIF is funded with €55 million in contributions from the EuroHPC JU, the German Federal Ministry of Research, Technology, and Space (BMBF), and the Ministry of Culture and Science of the state of North Rhine-Westphalia.

"With the JUPITER AI Factory, we are creating a powerful platform for AI that is leading the way for the European supercomputing infrastructure," said Prof. Dr. Dr. Thomas Lippert, Director of JSC. *Eric Gedenk*

PROJECT	JUPITER AI Factory (JAIF)
FUNDING AGENCY	JUPITER AI Factory (JAIF)
FUNDING AMOUNT	€55 million
RUNTIME	2025 – 2028
PARTNERS	Jülich Supercomputing Centre (JSC), the Center for Artificial Intelligence at RWTH Aachen University, the Fraunhofer Institutes for Applied Information Technology and for Intelligent and Information Systems, and the Hessian Center for Artificial Intelligence.

HammerHAI to Create an AI Factory for Science and Industry

As coordinator of one of the inaugural EuroHPC JU-funded AI Factories, HLRS will operate a new AI-optimized supercomputer and develop solutions to make AI more secure and accessible.

"The European and German governments have identified the creation of a robust, native European infrastructure for artificial intelligence as a high priority. HammerHAI will quickly help to address this need." Prof. Dr. Michael Resch, HLRS

The EuroHPC Joint Undertaking (EuroHPC JU) announced the creation of a new consortium called HammerHAI to establish Germany's first AI Factory. Coordinated by the High-Performance Computing Center of the University of Stuttgart (HLRS), HammerHAI will address the urgent demand for increased AI capabilities in academic research, start-ups, SMEs, and European industry, as well as in the public sector. The AI Factory will install a secure, AI-optimized supercomputing infrastructure, provide expert support services, and develop solutions to make it easier for researchers and companies of all sizes to access and use powerful AI technologies.

Bringing together prominent organizations within the German high-performance computing (HPC) and artificial intelligence communities, HammerHAI will quickly build an AI Factory on a solid foundation of existing infrastructure and expertise. Joining HLRS in the project is the Leibniz Supercomputing Centre, which like HLRS is also a member of the Gauss Centre for Supercomputing. In addition, the AI Factory will benefit from important contributions by the Gesellschaft für wissenschaftliche Datenverarbeitung Göttingen, the Karlsruhe Institute of Technology, and HPC/AI consultancy SICOS BW.

Building on HLRS's established focus on industry, engineering, and global challenges, HammerHAI will concentrate on addressing the needs of manufacturing, automotive and mobility, and start-ups and SMEs across all sectors, as well as scientific research.

The project budget of approximately €85 million is co-funded by the EuroHPC JU, the German Federal Ministry of Research, Technology, and Space, the Baden-Württemberg Ministry of Science, Research, and Art, the Bavarian Ministry of Science and the Arts, and the Lower Saxony Ministry for Science and Culture. HammerHAI was among the first seven European AI Factories announced by the EuroHPC JU. This was followed in March 2025 with the announcement of six additional European AI factories, including the JUPITER AI Factory (JAIF) based at the Jülich Supercomputing Centre. (For more information about JAIF, please visit page 25).

"The European and German governments have identified the creation of a robust, native European infrastructure for artificial intelligence as a high priority. HammerHAI will quickly help to address this need," said Prof. Dr. Michael Resch, Director of HLRS. "At the same time, the project will play an important role in implementing the development strategy at HLRS and the Gauss Centre for Supercomputing as we evolve to support the changing landscape of high-performance computing."

Breaking down barriers to AI

Currently, scientists and companies across Europe face a number of hurdles that have limited the adoption of artificial intelligence technologies. These include the prohibitive cost of installing AI computing hardware, the shortage of AI expertise, and data security concerns arising from Europe's current reliance on offshore cloud AI service providers.

By establishing a one-stop shop that addresses the needs of European AI users, HammerHAI aims to overcome these challenges. "Working closely with the European artificial intelligence community, HammerHAI will both deploy new AI-optimized computing infrastructure and develop a comprehensive portfolio of services and technologies," said HLRS Managing Director Dr. Bastian Koller, who will manage the project. "In this way, the AI factory will make it easier for scientists, startups, and SMEs to seize the opportunities that AI offers for discovery, innovation, automation, and decision making."

HammerHAI will build on the existing capabilities of Germany's national high-performance computing infrastructure and procure a new, large-scale AI-optimized supercomputer, which HLRS will operate on behalf of the EuroHPC Joint Undertaking and the project funders. The system will be tailored to the computational requirements of typical AI workloads in industry, including the training and usage of customized large language models, deep learning, and complex data analytics. The new system will also accommodate hybrid workflows that integrate HPC and AI, such as using HPC to generate synthetic datasets for the training of neural networks, or using AI

Members of the HammerHAI consortium met at HLRS in early 2025 to plan the start of the project.

PROJECT	HammerHAI
FUNDING AGENCY	EuroHPC Joint Undertaking German Federal Ministry of Research, Technology, and Space (BMBF) Baden-Württemberg Ministry of Science, Research, and Art Bavarian Ministry of Science and Art Lower Saxony Ministry for Science and Culture.
FUNDING AMOUNT	€ 85 millon
START DATE	April 1, 2025
PARTNERS	High-Performance Computing Center Stuttgart (coordinator) Leibniz Supercomputing Centre Gesellschaft für wissenschaftliche Datenverarbeitung Göttingen Karlsruhe Institute of Technology SICOS BW

to accelerate computationally demanding simulations through surrogate models. The hosting agreement for the new system with the EuroHPC JU was announced in March 2025.

The goal of HammerHAI is to provide a secure, accessible, and scalable resource that will become a cornerstone for Europe's AI ecosystem. It will support the development and testing of new applications as well as the migration and evolution of existing solutions that would benefit from the increased performance of an AI-optimized supercomputer. Migration will be facilitated through a cloud-native approach to HPC that offers web-based interfaces, an easy mechanism for data sharing, and the support of virtualization and containerization technologies. The system will also integrate pre-packaged AI frameworks like TensorFlow and PyTorch, and offer access to pre-trained AI models and large datasets.

Additionally, AI specialists at HammerHAI will promote the adoption of AI methods by assisting users as they migrate their applications from the cloud to HPC and offering comprehensive end-to-end support for the entire AI life cycle, from model development to the use of AI models in a production environment. Continuing professional education courses within the HammerHAI consortium will also empower application programmers to acquire the skills they need to fully exploit the potential of AI technologies.

AI made in Europe

Importantly, HammerHAI will provide a secure, Germany-based platform and data repository. It will place strict controls on access to the resulting AI models, the data on which they are based, and the query interfaces through which they are utilized. This will make it easier for European researchers, startups, and SMEs to adhere to compliance requirements concerning data security and the ethical use of AI. A publicly funded national high-performance computing center, HLRS is certified for information security under the ISO 27001 standard and has completed a Level 3 audit under the Trusted Information Security Assessment Exchange (TISAX), a framework for data centers with "very high protection needs."

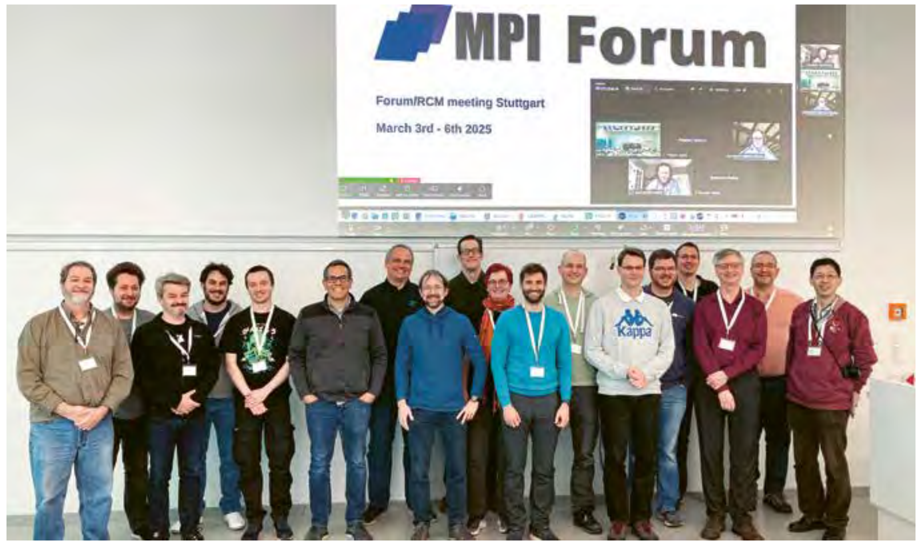
HammerHAI will be embedded within the German and European AI ecosystems. The project partners are closely integrated into numerous regional and national AI initiatives across Germany, and as coordinator of European HPC networking projects such as EuroCC 2, CASTIEL 2, and FFplus, HLRS will promote the availability of and access to the AI resources that HammerHAI will offer. In this way, the German AI Factory will become an important component of a Europe-wide effort to transform industry and support the public and academic sectors in the AI era. *Christopher Williams*

MPI Forum meets at HLRS on Path to MPI 5.0

The new standard for the Message Passing Interface will improve portability of applications across computing platforms and in containerized environments by introducing an application binary interface.

On March 3-6, 2025, members of the global community responsible for the continuing development of the Message Passing Interface (MPI) met at the High-Performance Computing Center Stuttgart (HLRS) to make final preparations for the next release of this critical parallel programming standard. Following a systematic and meticulous review and discussion of all facets of MPI, the results of the meeting will become the basis for the new MPI 5.0 standard.

First released in 1994, the MPI standard is one of the most important parallel programming models for HPC applications worldwide. It provides ways to exploit the most essential feature of high-performance computing using supercomputers: the ability to break a complex problem into many smaller parts and to run these processes across large numbers of processors in parallel. Structuring and managing parallel applications is extremely complex, as many processes must be connected and integrated in a way that uses a supercomputer's capabilities as efficiently as possible. The Message Passing Interface is a standard that enables programmers to synchronize these numerous processes and to manage communication among them. MPI has been widely adopted among programmers of high-performance computers worldwide because of its usability on most HPC hardware platforms.



MPI Forum members met in the Rühle-Saal at HLRS to prepare the new standard for parallel programming.

Bringing together vendors, researchers, and developers from the United States and Europe, the MPI Forum formed in the early 1990s to establish a standard that could address the need for a portable and performant way to program distributed memory systems. The upcoming MPI 5.0 standard is the first major update since MPI 4.0 was released in 2021. Each new standard implements changes to address the continually evolving demands that programmers face as a result of new computer hardware and new HPC applications.

The main feature of MPI 5.0 is the introduction of an application binary interface (ABI). This step addresses the fact that as the computing industry has grown in recent years, the diversity of new technologies has increased, raising problems with compatibility. The rise of artificial intelligence has also added to this complexity. Programmers must increasingly work in different environments and combine programs created using different technologies into workflows. The addition of an ABI to MPI 5.0 will enable interoperability of MPI libraries from different MPI vendors and on different HPC platforms. In addition, it will improve portability and performance for MPI-based applications in containerized environments.

HLRS is a longtime member of the MPI Forum, having first contributed to the creation of the MPI 2.0 standard, released in 1996. Currently, Dr. Christoph Niethammer and Dr. Tobias Haas represent HLRS within the MPI Forum as chapter chairs and as members of multiple working groups. Collaborating with other members of the MPI Forum, they participate in technical aspects of MPI software implementation, gather user perspectives to identify technical problems and understand user needs, and design course content for the instruction of parallel programming using MPI.

"In a world that increasingly relies on large-scale simulations and complex AI models, parallel programming is essential," Niethammer said. "For more than 30 years, the MPI standard has been a cornerstone of parallel programming, and is only growing in importance with the rise of artificial intelligence. The MPI Forum plays a key role in leading the continuing development of the MPI standard, and its next version will enable the computing community in science and industry to address the growing demands of modern applications."

Following the successful MPI Forum meeting in Stuttgart, a final ratification meeting will take place in June 2025. Once ratified, MPI 5.0 will be officially published. In the meantime, the MPI Forum has already started working on future improvements to the Message Passing Interface.
Christopher Williams

> **"In a world that increasingly relies on large-scale simulations and complex AI models, parallel programming is essential."**
> **Dr. Christoph Niethammer, HLRS**

EVENTS

JSC Hosts 12th NIC Symposium

Biannual review event brings together JSC users and staff for a multi-day interactive conference of presentations and discussions.

The John von Neumann Institute for Computing (NIC) supports research projects across a wide range of scientific disciplines, including physics, chemistry, life sciences, environmental sciences, AI and computer science, and applied mathematics.

On March 6–7, 2025, computational scientists presented their research results at the 12th NIC Symposium in Jülich. Originally held every two years, the event was postponed from 2022 to 2023 due to the COVID-19 pandemic and resumed in 2025. Now, the symposium is back on its biennial schedule. About 25 years after the first NIC Symposium took place in December 2001, this series of biannual meetings has established a valued tradition, showcasing some of the most advanced computational science at NIC. This year, the symposium reached a record attendance with over 220 participants.

Prof. Laurens Kuipers, better known as Kobus Kuipers, has been a new member of the Board of Directors of Forschungszentrum Jülich since March 2025 and welcomed all symposium participants at the beginning of the event. In his address, Prof. Kuipers emphasized NIC's long tradition of supporting computational science in Germany. Following this, the Director of the Jülich Supercomputing Centre (JSC) and NIC, Prof. Thomas Lippert, highlighted the enormous importance of AI and the role of JUPITER, Europe's first exascale supercomputer, in this context. JUPITER is scheduled to go online for production runs this summer.

Researchers from across the scientific spectrum presented their results and findings in numerous detailed talks throughout the event, as well as through an impressive 140 posters. The poster session not only served as a fertile ground for discussions among researchers but also provided an important platform for young scientists to present their work, gain visibility, and engage with experienced experts.

The 12th NIC Symposium had a lively poster session that brought together seasoned principal investigators with young researchers eager to present work done on JSC computing resources

It offered ample time and space to exchange ideas and experiences on JSC resources in an interdisciplinary setting.

As per tradition, researchers' recent activities and results were compiled in the NIC proceedings. The symposium and proceedings aimed at both computational scientists and practitioners, as well as the general public interested in the advancement of computational science and its applications in diverse contemporary research fields.
Alexander Trautmann

All accompanying materials, including the program, talks, posters, proceedings, and photographs, are available at: https://www.john-von-neumann-institut.de/en/news/nic-symposium/nic-symposium-2025

Open Lab Day at LRZ Brings Together Students and Staff for Visualization Showcases.

For the first time since the COVID-19 pandemic, staff at the Leibniz Supercomputing Centre welcomed students to its Open Lab Day. During the event, students presented visualization projects and participants engaged in conversation about the role of visualization in research.

In December, 2024, staff members at the Leibniz Supercomputing Centre (LRZ) hosted its Open Lab Day for the first time since the COVID-19 pandemic placed a multi-year moratorium on such in-person meetings. The event not only gives participants a view into LRZ's computing infrastructure, but it also invites students to showcase relevant projects.

As part of this year's event theme, "Re-Image," students were invited to present virtual reality (VR) games and scientific visualizations while also sharing their experiences in working with research data to create visualizations.

"Students of media informatics at the Ludwigs-Maximilians Universität München (LMU) and graduates of our "Virtual Reality" internship, which took place in the summer semester of 2024, presented projects and final theses," said Elizabeth Mayer, VR researcher at LRZ. "Internship participants were able to reinvent existing games or ideas. A total of 13 works were on display at the Open Lab Day, including eight final theses from the internship and two master's theses."

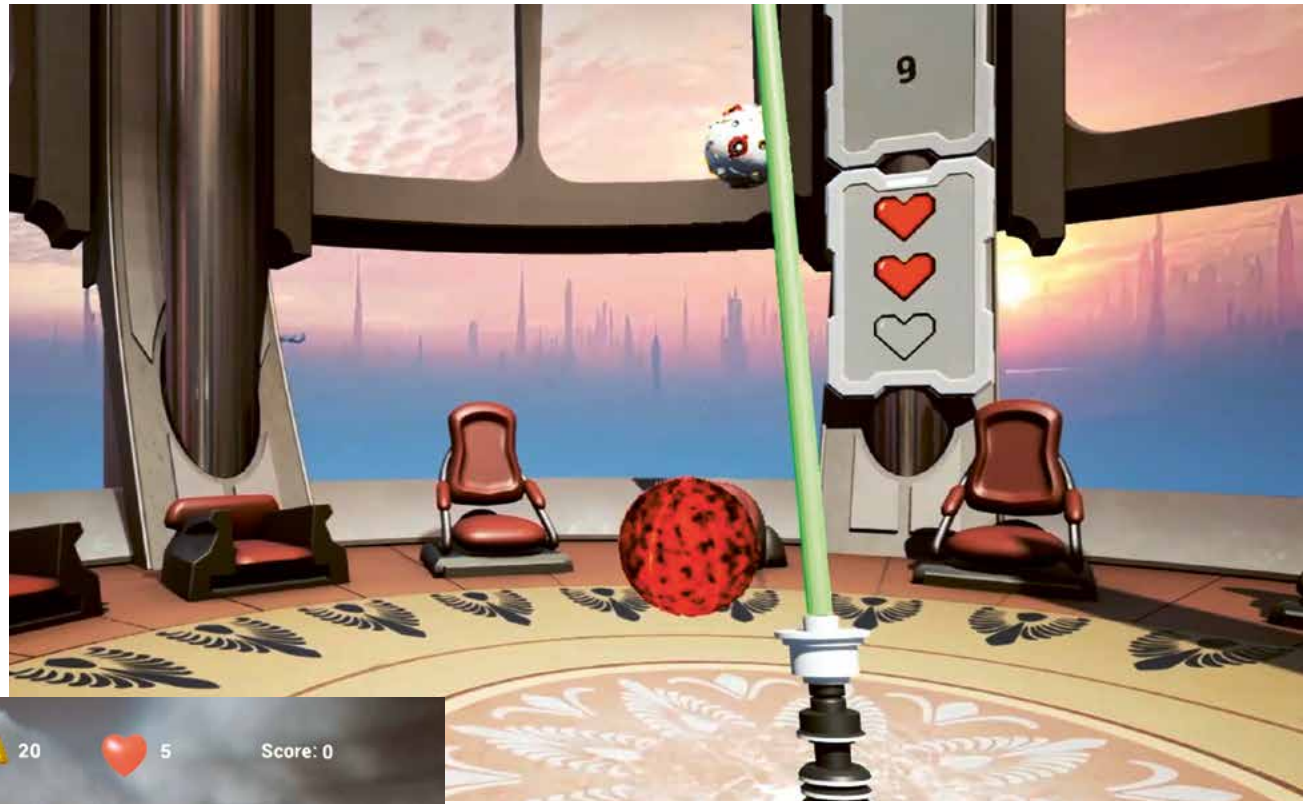
Dr. Thomas Odaker, Head of LRZ's Centre for VR and Visualization, or V2C, noted while the event brings different views together, showcasing student work was the central theme of the event. "This Open Lab Day was all about the students," he said. "We kept a low profile, but we enjoyed exchanging ideas and talking shop with our guests."

The event opened with Prof. Dieter Kranzlmüller, Chairman of the Board of Directors at LRZ, welcoming participants, followed by an introduction to the event by Odaker. Afterwards, Odaker and Mayer gave a presentation about the role of V2C in education at LRZ, and showcased the virtual reality summer internship program.

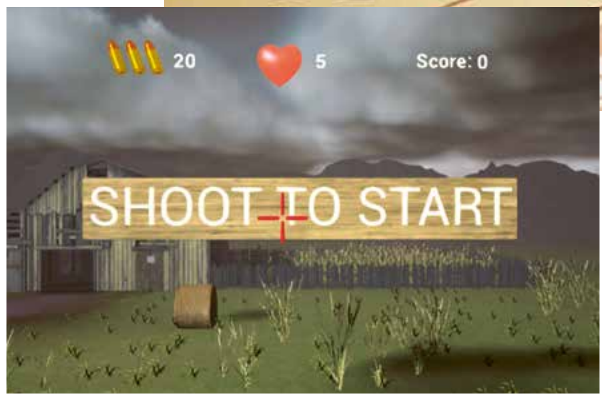
Participating students then were invited to give a "lightning talk" round to showcase their individual projects. Afterwards, event organizers opened the floor for presentation blocks that allowed students to showcase their visualization projects. Participants could also sign up for tours of the LRZ computer room while waiting for their turn to experience the VR showcases first-hand.

For Mayer, working with motivated students in V2C not only inspires and benefits the next generation of researchers, but also supports the core mission of LRZ visualization staff. "We usually have student assistants who help us with the various projects," she said. "The tasks range from preparing data sets and 3D modelling to transferring the data to LRZ's CAVE or the Powerwall at the center. We often find interested, motivated students among the trainees. LRZ is a digitization partner of the two Munich universities, and both of those institutions offer their own internships in the field of visualization as well. Bringing students into the CAVE is a core part of these courses' programs, and we can really build excitement during those visits and attract the next generation of motivated student interns." *Eric Gedenk*

"Bringing students into the CAVE is a core part of these courses."
Elizabeth Mayer, LRZ



Test your skills with the Photonsaber against a training drone in "Photonsaber Training VR" by Ludwig Degenhardt and Linus Stetter.



In "VR Chicken Shooter" Kristin Wloka and Yannick Erpelding created an experience for the CAVE inspired by the classic game "Moorhuhn".

The evil astronauts are invading! Defend your home planet against the invaders from earth in "Guardians of Uvadus" by Ehbal Ablimit, Ilinur Alimu, and Thu Mai Nguyen

News Briefs

JSC

Forschungszentrum Jülich Strengthens Quantum Research with Purchase of D-Wave Annealing Quantum Computer

JSC is permanently incorporating the D-Wave™ annealing quantum computer after three years of successful hosting and usage. It will remain a key part of the quantum computing infrastructure at Forschungszentrum Jülich and is expected to become increasingly important. The system will be connected to Europe's first and only Exascale system, JUPITER, to faciliate breakthroughs in artificial intelligence and quantum optimization applications.

WeatherGenerator Project Launches to Advance Earth System Science with AI

JSC is part of a newly formed consortium aiming to create a cutting-edge European foundation model for Earth Science – with the help of state-of-the-art AI and machine learning. The Weather Generator project officially launched in February, 2025 with a virtual kickoff meeting. The project will be carried out by 16 leading European institutions, including universities, research centers, meteorological institutes, and energy organizations. The project is coordinated by the European Centre for Medium-Range Weather Forecasts (ECMWF). The project aims to significantly enhance weather and climate predictions, optimize renewable energy planning, and provide better tools for mitigating the effects of extreme weather events, such as floods and heatwaves. By combining advanced machine learning techniques with high-performance computing and exploiting the new exascale computing capacity of JUPITER, the project will offer innovative solutions for both scientific research and real-world applications, benefiting sectors like renewable energy, disaster management, and public health.

JSC Coordinates SEANERGYS Project

The growing demand for high-end computing, particularly for artificial intelligence (AI), is driving growth in the number and size of supercomputers, with a commensurate increase in overall energy consumption for HPC and AI centers. To address this challenge, GCS centers focus on operating as efficiently as possible. To support this objective, the SEANERGYS project (Software for Efficient and Energy-Aware Supercomputers) will create an integrated European software solution for energy-efficient operation of HPC and AI systems that reduces the energy required to run today's and tomorrow's real-world applications and optimizes resource utilization and response times. These changes enable HPC/AI centers to reduce their operating costs and carbon footprint or generate more results for a given energy budget. The SEANERGYS software suite will include a Comprehensive Monitoring Infrastructure (CMI), an artificial intelligence data analytics system (AIDAS), and a Dynamic Scheduling and Resource Management System (DSRM). JSC coordinates the project with 15 other European academic and industrial partners. SEANERGYS will run from April 2025 through March 2029. The project builds upon the results of previous activities, like the DEEP and SEA projects, as well as the REGALE project. SEANERGYS is funded by the EuroHPC Joint Undertaking and the German Federal Ministry of Research, Technology, and Space.

For more information, please visit: https://www.fz-juelich.de/en/ias/jsc/projects/seanergys

JSC Bolsters European Sovereignty with DARE SGA1 Project

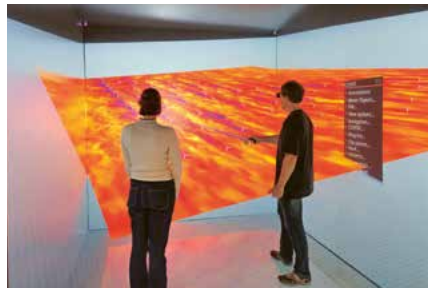
The Jülich Supercomputing Centre (JSC), together with Barcelona Supercomputing Centre (BSC), leads the software and applications activities in the DARE (Digital Autonomy with RISC-V in Europe) SGA1 project, which aims to enhance Europe's technological sovereignty in high-performance computing (HPC) and artificial intelligence (AI). DARE SGA1, which runs March 2025–February 2028, brings together 38 European partners from academia and industry. It focuses on developing next-generation European processors and computing systems alongside an optimized software ecosystem tailored for research and industrial applications. DARE SGA1 is the first phase of the development of high-TRL (Technical Readiness Level) RISC-V platforms for HPC and AI, which is coordinated by the DARE FPA (framework program agreement). A DARE SGA2 project is expected to evolve the prototypes created by DARE SAG1 to production readiness. Europe has been dependent on non-European hardware and software for its supercomputing infrastructure for a long time, which poses significant risks to security, economic stability, and technological competitiveness. DARE sets out to change this by utilizing the open RISC-V ecosystem and cutting-edge chiplet technology, paving the way for genuinely European products that will drive the continent's future supercomputing capabilities. In addition to technology development, the project defines the roadmap for Europe's post-exascale supercomputers. JSC, together with BSC, leads the software and applications technical area, which develops a competitive high-performance software stack and drives co-design with key application use-cases for the software stack and for the DARE-developed processors. The project is funded by the EuroHPC Joint Undertaking.

For more information, please visit: https://dare-riscv.eu/home/

HLRS

High-Level Ministers Visit HLRS

In March, HLRS welcomed Winfried Kretschmann (Minister-President of the State of Baden-Württemberg), Cem Özdemir (German Minister for Education and Research), and Petra Olschowski (Baden-Württemberg Minister for Science, Research, and Arts), providing a briefing on some of its latest activities and infrastructure improvements. In focus was HLRS's coordination of a new EuroHPC Joint Undertaking AI Factory called HammerHAI, as well as its role in operating computing infrastructure that supports German science and industry. (for more information on HammerHAI, please visit page 26). Commenting on the new AI Factory, Minister-President Kretschmann said, "Baden-Württemberg is at the forefront in the fields of supercomputing and AI, and is internationally competitive. HLRS plays a key role in this. The new Hunter supercomputer and the European AI Factory HammerHAI are groundbreaking milestones for our high-tech region, our economy, and our technological sovereignty." Minister Özdemir added, "HLRS and the AI Factory being established here with other partners are important pillars of the German and European AI ecosystem," while Science Minister Olschowski said, "Hunter ensures that research and development happens at the highest level, strengthening our ability to innovate and our economy."



Data Visualization Software Enhances Wind Energy Research

MERIDIONAL, an EU-funded project on wind energy research, is leveraging the power of VISTLE, an open-source scientific visualization software program developed by the High-Performance Computing Center Stuttgart. VISTLE enables interactive, high-fidelity visualizations of complex wind flow simulations, producing deeper insights into wind farm performance and aerodynamics. First released in 2012, VISTLE is a modular and extensible program designed to handle large-scale scientific datasets and a wide range of simulation formats. In its application to MERIDIONAL, the software facilitates real-time interaction with data to explore wind behavior in immersive virtual environments. Its remote rendering capabilities enable computationally intensive operations to be performed on high-performance clusters while maintaining seamless user interaction on local devices. VISTLE was used to create a remote, interactive visualization of the airflow around wind farms. A prototype developed in collaboration with energy company EnBW focuses on the He Dreiht wind farm in the North Sea, integrating a digital twin of the site with advanced simulation data. This visualization allows researchers to analyze both steady and unsteady airflows, helping to optimize wind turbine placement and performance. Furthermore, VISTLE's integration with immersive technologies, such as virtual and augmented reality, provides an intuitive platform for engineers, researchers, and stakeholders to interact with complex datasets.

HLRS Conference Explores Challenges of Modelling for Public Policy

The Department of Philosophy of Computational Sciences at the High-Performance Computing Center Stuttgart organized and hosted an international conference titled *Modeling for Policy*. Bringing together philosophers, social scientists, historians of science, and users of computing technologies for modeling, the event offered the opportunity to reflect on the practice of modeling and the challenges it can face in the context of public policy making. This latest event in the conference series *Science and Art of Simulation* covered a wide range of domains in which models are used, including public health, energy transformation, climate modeling, employment programs, and value chain analysis, among others. Within these contexts the event sought to define the capabilities and limits of simulation more clearly, explore how trust in models could be improved among policy makers and the general public, and make recommendations that could lead to better interactions between simulation scientists and society at large. Considering that modeling for public policy is embedded in a complex space influenced by constellations of social, political, economic, and cultural factors, the conference showed why a more complete understanding of the relationships among these domains will be needed to maintain public support of good science and to ensure that it continues to drive public policy.

HLRS Hosts Four S+T+ARTS Ec(h) of Artist Residencies

An initiative of the European Commission, S+T+ARTS is an international program that supports collaborations between artists, scientists, engineers, and researchers. In January it announced ten new artist residences under the S+T+ARTS EC(H)O residencies program. S+T+ARTS EC(H)O aims to promote ecologically conscious and human-compatible digital technology by supporting artists in developing innovative solutions that blend art, science, and technology. The artists' projects at HLRS are developing digital applications for exploring European cultural heritage, technological solutions inspired by nature, immersive visual narratives based on data, and how digital twins can bring sociocultural and ecological dimensions of urban environments to life. The Media Solution Center Baden-Württemberg (MSC) is a partner in the S+T+ARTS EC(H)O residency program. Founded by HLRS, the ZKM, and the Hochschule der Medien in 2018, the MSC facilitates collaboration among artists, scientists, and experts in high-performance computing. It has become an important node within European culture and creative industry networks, giving HLRS a role in EU initiatives promoting digital innovation in the arts.

LRZ

Welcome CoolMUC-4

The LRZ computer room got a new addition in December, 2024, welcoming CoolMUC-4 into the fold. The new Linux cluster initially offers some 12,000 cores based on Intel® Xeon®Platinum 8480+ processors and will be supplemented by a GPU partition later in 2025. After nine years of continuous operation, hundreds of thousands of calculations, and billions of core hours compute time, CoolMUC-2 and CoolMUC-3 retired in the middle of December.

For more information and system documentation, please visit: https://doku.lrz.de/coolmuc-4-1082337877.html

Kick-off for BayernKI

With the new BayernKI AI infrastructure, the Free State of Bavaria is accelerating the development and research of innovative AI methods and applications. The central AI infrastructure housed at the LRZ in Garching and the NHR@FAU provides dedicated support for academic AI research. The resources are dedicated to scientists at Bavarian universities and universities of applied sciences. In addition to having access to state-of-the-art hardware, users also receive individual support and consulting as well as access to a comprehensive course program. The new BayernKI portal gives researchers quick and easy access to the AI computing capacities at both locations. In the first stage, 320 H100 GPUs from NVIDIA are available as part of BayernKI. The BayernKI infrastructure at LRZ and NHR@FAU will be further expanded by 2028. Up to €55 million from the Hightech Agenda Bayern are available for BayernKI.

For more information, please visit: https://ki-in-bayern.de

LRZ at the EuroHPC Summit 2025

In March 2025, the EuroHPC JU gathered scientists and managers at the ICE Center in Krakow to discuss developments in high-performance computing (HPC). The two focus topics in 2025: Quantum Computing and Artificial Intelligence (AI). Several LRZ colleagues presented their own solutions and lectures - in the field of AI, for example, as a consortium member of the German AI factory HammerHAI. LRZ specialists also presented Euro-Q-Exa, the hybrid quantum supercomputer based on superconducting technology. This will be hosted at the LRZ for the EuroHPC JU and is due to go into operation later this year.

From more information, please visit: https://www.lrz.de/news/detail/das-lrz-auf-dem-eurohpc-summit-2025

New Industry Advisory Board at LRZ

The Leibniz Supercomputing Centre (LRZ) has established an Industry Advisory Board comprising three distinguished IT experts – Wolf I. Faecks (consultant), Dr. Wieland Holfelder (VP Engineering Centre Google Munich), and Stefan Wagner (MD SAP Labs Munich) – to enhance collaboration with industry and align its services with emerging technological and economic trends. This initiative aims to bridge the gap between scientific research and industrial applications, particularly in the evolving landscape of artificial intelligence and digital transformation. The Industry Advisory Board will discuss technological advancements, workforce development, and strategies for efficiently commercializing research outcomes with LRZ leadership.

For more information, please visit: https://www.lrz.de/en/news/detail/new-advisory-board-with-industry-experts-at-lrz

LRZ puts HPC System CoolMUC-4 Into Operation

A new HPC cluster has gone into operation at the LRZ. CoolMUC combines the performance of various processor architectures and is now available to researchers at Bavarian Bavarian universities with immediate effect. A total of 119computing nodes - offer concentrated performance, including 106 with Intel Sapphire Rapids chips, which are equipped with 512 Giga-bytes of memory and 112 computing cores, particularly for parallel applications with low memory requirements are particularly suitable. For more data-intensive jobs, the 12 Ice Lake nodes are used for more data-intensive jobs, each with 80 computing cores cores and one terabyte of short-term memory. The system is supplemented the system with "Teramem", a node with a Cooper Lake processor, 96 computing cores and around 6 terabytes of of RAM. The current CoolMUC will soon be with GPUs and will be expanded with additional CPUs by 2026. In addition to the classic HPC disciplines, more and more disciplines such as biology, environmental and life sciences life sciences or social and economic sciences sciences are using the computing cluster for their research.

For more information, please visit: tiny.badw.de/XH6Y1.

Fighting Alzheimer's with HPC

CLARA is an interdisciplinary European Centre of Excellence that uses computing power to fight neurodegenerative diseases such as Alzheimer's and Parkinson's. The LRZ is involved in this endeavor in supporting researchers who wish to team up traditional HPC with AI and quantum computing to better categorize neurodegenerative diesases' symptoms.

For more information, please visit: https://www.it4i.cz/en/about/infoservice/press-releases/the-new-clara-research-centre-for-neurodegenerative-disease-research

“After two years of traditional software consulting, I knew I wanted to explore new areas.” Fritz Niesel, JSC

Fritz Niesel has a passion for engaging with industry researchers to develop AI-enabled solutions for their businesses. After several years in the private sector, Niesel joined JSC in 2022 to help spearhead the center's efforts to bolster its AI support for academia and industry alike.

As the JUPITER Era Arrives, Early Career Staffer Supports AI Uptake in Industry

Fritz Niesel, JSC

Starting his secondary education at Baden-Württemberg Cooperative University (DHBW) in 2014, Fritz Niesel knew that he wanted to apply his interest in math toward practical applications. He studied business IT at DHBW, and as part of his work-study arrangement, worked at German software giant SAP. In that role, he worked as a business process consultant, supporting organizations interested in integrating SAP software in their workflows and developing customized solutions of the standard software for their specific needs.

“After two years of traditional software consulting, I knew I wanted to explore new areas,” Niesel said. “I was particularly drawn to the rising popularity of artificial intelligence (AI), and that led me to pursue my master's degree in data science at the University of Mannheim.”

Shortly after completing his graduate studies in 2022, Niesel joined the Jülich Supercomputing Centre (JSC) as an AI consultant, and he couldn't be happier with the work. “Ultimately, my main passion is problem solving, engaging with customers, and developing AI-driven solutions that can help them address unique challenges they face in their business,” he said.

Niesel's passions align closely with JSC's initiatives to increase the use of HPC and AI workflows for industry, with a specific focus on supporting small and medium enterprises who may not have as much experience using large-scale computational tools to accelerate research and development.

To that end, JSC recently partnered with several universities in North Rhine-Westphalia as well as multiple Fraunhofer institutes to start the AI service center, WestAI. The initiative offers industry researchers at all experience levels an opportunity to partner with consultants to talk through their challenges, participate in AI-focused training courses, access AI-ready high-performance computing (HPC) resources, and research collaboratively with WestAI participating organizations.

As one of the leads on WestAI's Development Service and AI Consulting team, Niesel is working with companies ranging from engineering firms to educational tool suppliers. He recently collaborated with e-learning platform provider u-form GmbH to help the organization leverage the power of large language models (LLMs). Essentially, Niesel worked with the organization to train several commercial and open-source LLMs using German-language, essay-style exam questions and evaluate how well these different applications could grade open-ended questions. “I really liked working on this project,” Niesel said. “Even though this was not a topic I studied, and LLMs were not even in the public consciousness during my studies, I learned a lot from this experience. The company hoped to use an open-source solution for this work, but we ultimately found that a proprietary, commercial LLM was better suited for the task.”

Niesel noted that his efforts to work closer with non-traditional users at HPC centers falls in line with the larger mission of public HPC centers – working with a wide

INSIDE
GCS

“On some level, all kinds of AI problems are comparable: you need to have some input data and have a vision for what you need as output data.”
Fritz Niesel, JSC

range of subjects and skill levels to support societally beneficial research. By working closely with a diverse range of subject-matter experts at JSC, Niesel feels confident that JSC can offer scientists and engineers from across the research spectrum ways to successfully integrate AI into their research. “I’m confident to say that I can add value to almost any kind of entry level problem through consulting with researchers,” he said. “On some level, all kinds of AI problems are comparable: you need to have some input data and have a vision for what you need as output data. In every case, we evaluate if researchers goals are feasible and then support them with the resources they need to achieve those goals.”

In 2025, Niesel and his JSC colleagues are also beginning the JUPITER AI Factory (JAIF), a project jointly funded by the EuroHPC Joint Undertaking, the German Federal Ministry of Research, Technology, and Space, the Ministry of Culture and Science of the state of North Rhine-Westphalia, and the Hessian Ministry of Science and Research, Arts and Culture (for more information on JAIF, visit page 25). Together with HammerHAI, an AI Factory being led by the High-Performance Computing Center Stuttgart with participation from the Leibniz Supercomputing Centre, all GCS centers will offer increased support for AI projects and researchers interested in including AI workflows in their research. (For more information on HammerHAI, visit page 26).

Niesel pointed out that as part of the Gauss Centre for Supercomputing, JSC staff working on JAIF have a robust network to share best practices on AI topics ranging from optimizing AI workflows on specific hardware to ensuring compliance with legal and security requirements for projects coming from industry. As the AI Factories bring in more diverse research teams, he indicated that such knowledge sharing will become increasingly important.

While still relatively early in his career, Niesel has carved out a role in JSC that not only scratches his itch for problem-solving and collaboration, but also allows him to experiment with different paths to focus on throughout his career. His AI consulting work demonstrated to him that he has a passion for communicating, bringing diverse perspectives together, and breaking down complex AI topics to help researchers craft appropriate solutions. “In the years to come, I can envision leading a dedicated “AI-for-industry” group that is purely focused on bridging the gap between consulting and implementation,” he said. Working at a place like JSC – where the needed hardware, infrastructure, expertise, and passion come together – allows Niesel to dream big about the future of AI and about his career in the years to come.
Eric Gedenk

With JUPITER’s arrival this year, Niesel and his JSC colleagues will have a powerful new tool to support AI research in Jülich.

The HUNTER supercomputer at HLRS.

High-Performance Computing Center Stuttgart

The High-Performance Computing Center Stuttgart (HLRS) was established in 1996 as Germany's first national high-performance computing center. A research institution affiliated with both GCS and the University of Stuttgart, HLRS provides infrastructure and services for HPC, data analytics, visualization, and artificial intelligence to academic users and industry across many scientific disciplines, with an emphasis on computational engineering and applied science.

Supercomputing for industry

Through a public-private joint venture called hww (Höchstleistungsrechner für Wissenschaft und Wirtschaft), HLRS ensures that industry always has access to state-of-the-art HPC technologies. HLRS also helped to found SICOS BW GmbH, which assists small and medium-sized enterprises in accessing HPC technologies and resources. Additionally, HLRS cofounded the Supercomputing-Akademie, a training program that addresses the unique needs of industrial HPC users.

Guiding the future of supercomputing

HLRS scientists participate in dozens of funded research projects, working closely with academic and industrial partners to address key problems facing the future of computing. Projects develop new technologies and address global challenges where supercomputing can provide practical solutions. With the support of the EuroHPC Joint Undertaking, HLRS is also currently coordinating efforts to build and integrate HPC competencies across Europe. The center is certified for environmental management under the EU's Eco-Management and Audit Scheme (EMAS) and for information security under the ISO 270001 standard.

H L R S
High-Performance Computing Center Stuttgart

CONTACT

High-Performance Computing Center Stuttgart (HLRS), University of Stuttgart

Prof. Dr.-Ing. Dr. h.c. Dr. h.c. Hon.-Prof. Michael M. Resch
Nobelstraße 19, 70569 Stuttgart, Germany
Phone +49 - 711 - 685 - 8 72 69
resch@hlrs.de
www.hlrs.de

Compute servers currently operated by HLRS

System	Size	Peak Performance (Tflop/s)	Purpose	User Community
HPE CRAY EX4000 (HUNTER)	APU: AMD Instinct MI300A, 188 nodes. 512 GB HBM3 (~5.3 TB/s) memory CPU: AMD EPYC 9374F, 256 nodes. 768 GB DDR5-4800 memory Cray ClustorStor E2000, 25 PB on 2,120 disks HPE Slingshot 11 Dragonfly (APU: 4 x 200 Gbps per node)	48.1 PFlop	Capability Computing	German and European research organizations and industry
HPE APOLLO 6500 GEN10 PLUS (HAWK AI EXPANSION)	24 nodes, 192 NVIDIA A100 GPUs	120 Pflops AI performance	Machine learning and artificial intelligence applications	German and European research organizations and industry
NEC CLUSTER (VULCAN, VULCAN 2)	CPU Intel 6230 (8 nodes), Intel 6258 (96 nodes), Intel 6238 (72 nodes), Intel 4112 (9 nodes), AMD 9124 (24 nodes), AMD 9334 (154 nodes), AMD 7302 (3 nodes), AMD 7642 (4 nodes)	CPU performance: 802 TF	Capacity Computing	German Universities, Research Institutes, and Industry
	GPU NVIDIA A30 (24 nodes), AMD MI50 (4 nodes, 8 GPUs per node), NVIDIA RTX4000 (3 nodes), AMD WX8200 (6 nodes)	GPU performance Single precision: 763 TF Single precision: 763 TF		

Modular Supercomputer JUWELS at the Jülich Supercomputing Centre.

Jülich Supercomputing Centre Forschungszentrum Jülich

The Jülich Supercomputing Centre (JSC) at Forschungszentrum Jülich is committed to enabling scientists and engineers to explore some of the most complex grand challenges facing science and society. Our research is performed through collaborative infrastructures, exploiting extreme-scale supercomputing, AI at scale, quantum computing, and federated data services.

Provision of supercomputer resources:

JSC provides access to supercomputing resources of the highest performance for research projects coming from academia, research organizations, and industry. Users gain access for projects across the science and engineering spectrum in the fields of modeling and computer science.

Core tasks of JSC are:

- Supercomputer-oriented research and development in selected fields of physics and other natural sciences by research groups and in technology, e.g. by doing co-design together with leading HPC companies.
- Implementation of strategic support infrastructures including community-oriented simulation and data laboratories and cross-sectional teams, e.g. on mathematical methods and algorithms and parallel performance tools, enabling the effective usage of the supercomputer resources.
- Cutting-edge quantum computing research and access through the Jülich UNified Infrastructure for Quantum computing (JUNIQ).
- Higher education for master and doctoral students in close cooperation with neighbouring universities.

JÜLICH Forschungszentrum | JÜLICH SUPERCOMPUTING CENTRE

CONTACT

Jülich Supercomputing Centre (JSC)
Forschungszentrum Jülich

Prof. Dr. Dr. Thomas Lippert
Wilhelm-Johnen-Straße, 52425 Jülich, Germany
Phone +49 - 24 61- 61- 64 02
th.lippert@fz-juelich.de
www.fz-juelich.de/jsc

Compute servers currently operated by JSC

System	Size	Peak Performance (Tflop/s)	Purpose	User Community
MODULAR SUPERCOMPUTER "JUWELS"	Cluster (Atos): 10 cells, 2,567 nodes 122,768 cores Intel Skylake 224 NVIDIA V100 GPUs 275 TByte memory	12,266	Capability Computing	European and German Universities and Research Institutes
	Booster (Atos): 39 racks, 936 nodes 44,928 cores AMD EPYC Rome 3,744 NVIDIA A100 GPUs 629 TByte memory	75,020		
SUPERCOMPUTER "JURECA"	Data-Centric Cluster (Atos): 768 nodes, 98,304 cores AMD EPYC Rome 768 NVIDIA A100 GPUs 443 TByte memory	18,515	Capacity and Capability-Computing	European and German Universities, Research Institutes, and Industry
JUPITER DEVELOPMENT SYSTEM "JEDI"	1 rack, 48 nodes 13.824 cores 192 NVIDIA GH200 Superchips 40,5 TByte memory	10,264	Capability Computing	European and German Universities and Research Institutes
ATOS CLUSTER "JUSUF"	205 nodes, 26,240 cores AMD EPYC Rome 61 NVIDIA V100 GPUs 52 TByte memory	1,372	Capacity and Cloud Computing	European and German Universities and Research Institutes through Human Brain Project
MODULAR SUPERCOMPUTER "DEEP-EST" (PROTOTYPE)	Cluster: 50 nodes, 1,200 cores Intel Xeon Gold 6146 9.6 TByte memory + 25.6 TByte NVM	45	Capacity Computing (low-/medium-scalable code parts)	Partners of the "DEEP" and "SEA" EU-project series and interested users through Early Access Programme
	Booster: 75 nodes, 600 cores Intel Xeon Silver 4215 75 NVIDIA V100 GPUs 6 TByte memory 38 Tbyte NVM	549	Capacity and Capability Computing (high-scalable code parts)	
	Data Analytics Module: 16 nodes, 768 cores Intel Xeon Platinum 8260 16 NVIDIA V100 GPUs 6.7 TByte memory 43 TByte NVM	170	Capacity and Capability Computing (data analytics codes)	
D-WAVE QUANTUM ANNEALER "JUPSI"	More than 5,000 qubits	No classical performance measure applicable	Quantum Computing	German Universities and Research Institutes (10%) Industry Applications and D-Wave customers (90%)

The SuperMUC-NG supercomputer at LRZ.

Leibniz Supercomputing Centre

For more than six decades, the Leibniz Supercomputing Centre (Leibniz-Rechenzentrum, LRZ) has been at the forefront of its field as a world-class high-performance computing center dedicated to providing an optimal IT infrastructure to its clients throughout the scientific community – from students to postdocs to renowned scientists – and in a broad spectrum of disciplines – from astrophysics and engineering to life sciences and digital humanities.

Leadership in HPC and HPDA

Located on the research campus in Garching near Munich, the LRZ is a leadership-class HPC and HPDA facility delivering top-tier supercomputing resources and services on the national and European levels. Top-notch specialists for HPC code portability and scalability support the broad user base at LRZ and ensure that the systems are running their operations in the most energy efficient way possible.

Leibniz Supercomputing Centre
of the Bavarian Academy of Sciences and Humanities

Quantum and Future Computing at LRZ

LRZ is leading the way forward in the field of future computing, focusing on emerging technologies like quantum computing and integrating AI on large-scale HPC systems. A robust education program that touches on HPC, machine learning, artificial intelligence, and big data complements LRZ's offerings.

IT backbone for Bavarian science

In addition to its role as a national supercomputing center, LRZ is also the IT service provider for all Munich universities as well as research organizations throughout Bavaria.

CONTACT

Leibniz Supercomputing Centre (LRZ)

Prof. Dr. Dieter Kranzlmüller
Boltzmannstraße 1, 85748 Garching near Munich, Germany
Phone +49 - 89 - 358 - 31- 80 00
leitung@lrz.de
www.lrz.de

Compute servers currently operated by LRZ

System	Size	Peak Performance (Tflop/s)	Purpose	User Community
SUPERMUC-NG PHASE 1 INTEL/LENOVO THINKSYSTEM	6,336 direct hot-water cooled compute nodes, 304,128 cores, Intel Xeon Platinum 8174, 608 TByte of memory, Omni-Path 100G interconnect	26,300	Capability Computing	German universities and research institutes(Tier-1)
	144 direct hot-water cooled compute nodes, 6,912 cores, Intel Xeon Platinum 8174, 111 TByte of memory, Omni-Path 100G interconnect	600	Capability Computing	
SUPERMUC-NG COMPUTE CLOUD	64 air-cooled nodes, 5,120 cores, Intel Xeon Gold 6148, 64 Nvidia Tesla V100	644 (CPUs + GPUs) 7,168 AI Performance*	Cloud Computing	German Universities and Research Institutes(Tier-1)
SUPERMUC-NG PHASE 2	240 direct hot-water cooled compute nodes, 26.880 Intel Xeon Platinum 8480+ compute cores (Sapphire Rapids), 122,88 TByte of memory, 960 accelerators (Intel Ponte Vecchio), NVIDIA HDR Infiniband interconnect	Pilot operation started	Capability Computing & Machine Learning, AI applications	German Universities and Research Institutes(Tier-1)
COOLMUC (4th GENERATION)	119 direct hot-water cooled compute nodes, 12,928 cores Intel Xeon Platinum 8480+ (Sapphire Rapids) and others, 76 TByte memory, NVIDIA HDR Infiniband interconnect	807	Capability Computing	Bavarian Universities (Tier-2)
LRZ AI SYSTEMS	17 nodes (NVIDIA GPU-based), HDR Infinband 88 NVIDIA GPUs, 3,328 GB HBMemory 1,424 CPUs, 5,824 GB DDR4 Memory	2,302 66,690 AI Performance*	Machine Learning, AI applications	Bavarian Universities
CEREBRAS CS-2	1 node with 850,000 compute cores, 40GB SRAM, 20 PB/s memory bandwidth and 220Pb/s interconnect	3,570,000 AI Performance* (estimate based on arXiv:2204.03775)	Purpose-built Deep Learning System	Select users
LRZ QUANTUM COMPUTING RESOURCES	Eviden Qaptiva 1		Quantum simulation	Bavarian Universities
	Eviden Qaptiva 2		Quantum simulation	Select users
	QMWare		Quantum simulation	Bavarian Universities
	IQM 5-qubit system		Quantum computation	Select users
	IQM 20-qubit system 1		Quantum computation	Select users
	IQM 20-qubit system 2		Quantum computation	Select users
	AQT 20-qubit ion-trap system		Quantum computation	Select users

*AI Performance refers to GPU peak performance for FP16 operations. For Nvidia GPUs, it is specific to different architectures. P100 architecture: CUDA core performance. V100 architecture: Mixed precision Tensor Core performance. A100: Structured sparsity Tensor Core performance.

Would you like to keep up with the newest research and news happening at the GCS centers?

Please send your postal address to pr@hlrs.de and you will recieve InSiDE magazine twice a year.

IMPRINT

InSiDE is published two times a year by the Gauss Centre for Supercomputing e.V., Alexanderplatz 1, 10178 Berlin

Editor-in-Chief: Michael Resch, HLRS Höchstleistungs-rechenzentrum Stuttgart, Nobelstaße 19, 70569 Stuttgart, resch@hlrs.de

Executive Editor: Eric Gedenk, GCS, eric.gedenk@lrz.de

Art Direction and Design:
GROOTHUIS. Gesellschaft der Ideen und Passionen mbH für Kommunikation und Medien, Marketing und Gestaltung: Lucy Lepstück; www.groothuis.de

Print:
Gutenberg Beuys Feindruckerei GmbH

This magazine has been printed climate-neutral on paper that has been certified by FSC®.

RECYCLED Paper made from recycled material FSC® C009051

Druckprodukt mit finanziellem Klimabeitrag ClimatePartner.com/10951-2505-1002

Contributing authors: Eric Gedenk, GCS, eric.gedenk@lrz.de; Christopher Williams, HLRS, williams@hlrs.de; Susanne Vieser, LRZ, susanne.vieser@lrz.de; Benedikt von St. Vieth, b.von.st.vieth@fz-juelich.de; Maximillian Tandi, JSC, m.tandi@fz-juelich.de

Photo credits:
Cover page 1/2: unsplash.com/lekoarts; pg. 3: Alessandro Podo, Sascha Kreklau, HLRS; pg. 6: FZ Jülich/Sonja Habbinga; pg. 8: Forschungszentrum Jülich; pg. 11: LRZ; pg. 14: Rocco Gentile/JACS; pg. 16: generated with Firefly by Groothuis; pg. 19: Hannes Bergström; pg. 21/22: Katharina Steinegger; pg. 24: unsplash.com/Maxim Berg; pg. 29: HLRS; pg. 31: HLRS; pg. 33: Forschungszentrum Jülich/Jenö Gellinek; pg. 35: Ludwig Degenhardt, Linus Stetter, Kristin Wloka, Yannick Erpelding, Ehbal Ablimit, Ilinur Alimu, and Thu Mai Nguyen; pg. 36/37: Forschungzentrum Jülich, HLRS; pg. 38/39: HLRS, LRZ; pg. 40: Forschungszentrum Jülich/Sascha Kreklau; pg. 43: X Forschungszentrum Jülich; pg. 44: HLRS; pg. 46: Forschungszentrum Jülich; pg. 48: LRZ

InSiDE magazine (German: Innovatives Supercomputing in Deutschland) is the bi-annual publication of the Gauss Centre for Supercomputing, showcasing recent highlights and scientific accomplishments from users at Germany's three national supercomputing centers. GCS was founded in 2007 as a partnership between the High-Performance Computing Center Stuttgart, Jülich Supercomputing Centre, and the Leibniz Supercomputing Centre. It is jointly funded by the German Ministry of Education and Science (Bundesministerium für Bildung und Forschung – BMBF) and the corresponding ministries of the three states of Baden-Württemberg, North Rhine-Westphalia, and Bavaria.

GCS
Gauss Centre for Supercomputing

www.gauss-centre.eu

Funding for GCS HPC resources is provided by:

Federal Ministry of Research, Technology and Space

Bavarian State Ministry of Science and the Arts

Baden-Württemberg Ministry of Science, Research and Arts

Ministry of Culture and Science of the State of North Rhine-Westphalia