InSiDE

Innovatives Supercomputing in Deutschland

No 25-2 · Autumn 2025

JUPITER Rising

Local, national, and European leaders came together to celebrate the inauguration of Europe's first exascale supercomputer.

Record-Breaking Space Simulation

Researchers used LRZ HPC resources to run the largest-ever simulation of astrophysical turbulence and magnetism in the interstellar medium.

AI Factory Focus

EU Commissioner Henna Virkkunen visited HLRS and JSC on her tour of AI Factories co-funded by the EU and national authorities.

GCS
Gauss Centre for Supercomputing

In recent years, the explosion of artificial intelligence (AI) applications and the crossing of the exascale threshold have ushered in a new era of high-performance computing (HPC). For over a decade, AI and exascale served as the horizon, with our centers pooling resources, expertise, and systems to ensure our users have access to world-class research tools. Now that we have arrived, we are setting our sights on the next horizon – integrating these powerful tools into workflows that will benefit both academia and industry while we help lead European efforts to enhance technological sovereignty and next-generation, made-in-Europe products and tools.

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 celebrate reaching a new horizon at our institutions while also looking back at the steps that got us to this point.

Recently, the Jülich Supercomputing Centre welcomed Europe's first exascale supercomputer, JUPITER. The inauguration drew local, national, and European leaders, including Federal Chancellor Friedrich Merz, and highlighted the key role the system will play in advancing European HPC and AI excellence in the years to come (page 6). A core part of Europe's strategy for this new computing horizon comes through investments in AI Factories. All three GCS centers are participating in AI Factories, and HLRS and JSC are coordinating two of these large-scale investments. This year, Henna Virkkunen, the European Commission Executive Vice President for Tech Sovereignty, Security, and Democracy, visited both centers as part of her tour of AI Factories (page 12).

While we are excited about the future of HPC and AI with these investments, we also look back at all the innovation and effort that got us to this point. In celebrating its 25th anniversary as a national HPC center, LRZ looks back on the systems, technical innovations, and scientific output that helped pave the way for exascale computing and advanced AI (page 9). At the same time, LRZ and technology partner HPE recently deployed "Blue Cubs," a test system that helps prepare staff and researchers alike for the center's next flagship supercomputer, Blue Lion (page 25).

While our centers have long supported both academic and commercial research, we are increasing our outreach to researchers in industry. GCS and JSC recently hosted the second "GCS Industry Day," which brought together our computing experts with industry researchers to highlight how GCS HPC systems can benefit their work, ultimately strengthening German and European economic competitiveness in the process (page 34). HLRS also recently hosted a five-part series to explore how HPC-accelerated AI workflows can benefit small and medium-sized enterprises (page 31).

Our centers continue to situate themselves as leaders in this new computing horizon. We have heavily invested in a world-class, diverse mix of computing architectures, brought in a deep roster of experts in AI and domain sciences, and are actively supporting pan-European initiatives to bolster the continent's computing capabilities and economic competitiveness. We feel strongly that these are the pillars that best support this time of rapid change in our industry.

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

"We are witnessing a historic European pioneering project."
Federal Chancellor Friedrich Merz

Leaders from the local to the European level came together to launch the JUPITER supercomputer. From left: Prof. Dr. Dr. Thomas Lippert, Director of Jülich Supercomputing Centre (JSC); Prof. Dr. Kristel Michielsen, Director of Jülich Supercomputing Centre (JSC); Ekaterina Zaharieva, Commissioner Startups, Research and Innovation, European Commission; Hendrik Wüst, Minister-President of North Rhine-Westphalia; Prof. Dr. Astrid Lambrecht, Chair of the Board of Directors at FZJ; Friedrich Merz, German Federal Chancellor; Dorothee Bär, Federal Minister for Research, Technology and Space; Karsten Wildberger, Federal Ministry for Digital Transformation and Government Modernization; Ina Brandes, Minister for Culture and Science of North Rhine-Westphalia; Prof. Dr. Laurens Kuipers, Member of the Board of Directors at FZJ.

A Highly Anticipated Celebration: Supercomputer JUPITER Put Into Operation

On Sept. 5, 2025, German state, federal, and European dignitaries came together to officially welcome JUPITER, Europe's first exascale supercomputer. The system marks a major leap in supercomputing capabilities, with the promise of significant research advancements for academia and industry.

"My team, myself, and JUPITER are celebrating an extraordinary festive day today," said Jülich Supercomputing Centre (JSC) Director Thomas Lippert at the beginning of his speech on the September 5 inauguration of Europe's first exascale supercomputer, JUPITER.

Forschungszentrum Jülich (FZJ) looks different today, and it feels different too: there is a dignified, festive atmosphere in the air, and rightly so. Today, Jülich, North Rhine-Westphalia, Germany, and Europe are celebrating the launch of a "new dimension in computing." The event also gave German Chancellor Friedrich Merz his first look at the system, and during his remarks, he put the accomplishment into perspective. "You have paved the way for innovations that are only now becoming possible thanks to this supercomputer and that, for the most part, we probably cannot even imagine today," he said.

For the event, FZJ placed a huge marquee in the middle of its expansive campus – between the campus's "Seecasino" and the Jülich Supercomputing Centre (JSC) – which set the stage for the celebratory inauguration . Host Johannes Döbbelt guided the audience through the program with a pleasant ease.

In front of around 800 invited guests on-site in Jülich and numerous viewers from all over the world who joined the event online, Europe's fastest supercomputer and the fourth fastest in the world is symbolically put into operation. As the first of its kind, it will reach 1 exaFLOP/s – that is one billion billion, or one quintillion floating-point operations per second. Chair of the FZJ Board of Directors Astrid Lambrecht emphasized this point during the press briefing: "That's a 1 followed by 18 zeros. It's a very large number."

It's a day many have been waiting for: those who designed the supercomputer, those who built it or provided financial support, and those who use – or will use – it. The system was funded by the EuroHPC Joint Undertaking and the German Federal Ministry of Research, Technology, and Space and the North Rhein-Westphalia Ministry of Culture and Science through the Gauss Centre for Supercomputing (GCS).

In his remarks, Lippert made a point to put emphasis on the "we" in reference to JUPITER: It is intended to underscore the great significance of Europe's first exascale computer. JUPITER has already become an essential research tool for JSC, Forschungszentrum Jülich, North Rhine-Westphalia, Germany, and Europe. The system holds great promise in helping researchers develop European language models that can compete with those of the US, global climate forecasts with local precision, personalized models to better understand and treat serious diseases, and advance cosmology and physics models so that JUPITER could even "replay" the universe, among countless other applications.

"JUPITER is a cornerstone of Europe's digital transformation."
Henna Virkkunen, European Commission Executive Vice-President for Tech Sovereignty, Security and Democracy

European, national, and state decision makers converge to celebrate achievement

As a truly European achievement and demonstration of the power of European scientific investment, the JUPITER inauguration was attended by many high-ranking political figures.

Federal Chancellor Friedrich Merz opened his speech by putting the achievement in historical context. “Exactly 48 years ago to the day, the Voyager 1 space probe was launched from Cape Canaveral in the United States, under the eyes of the world, heading for Jupiter, he said. “JUPITER is a technological and scientific breakthrough and the starting signal for new scientific progress. It is a historic pioneering project. It is pure coincidence that we are inaugurating JUPITER in Jülich on this anniversary day, but it is a coincidence with great symbolic power.”

Following Merz’s speech, big words also came from Brussels. In her video message, Henna Virkkunen, the European Commission’s Executive Vice-President for Tech Sovereignty, Security and Democracy, emphasized that JUPITER stands as proof of Germany’s long-standing leadership in the field of high-performance computing. At the same time, she announced preparations for European-funded AI Gigafactories and new investments in AI. She described JUPITER as a key cornerstone of digital transformation. “This is not just the inauguration of a supercomputer. It is the beginning of a new era of European excellence in high-performance computing, artificial intelligence, and science,” she said.

Hendrik Wüst, Minister President of North Rhine-Westphalia, followed Virkkunen, bringing a touch of Rhenish pragmatism to the stage. In his words of thanks to the funding partners, he emphasized that JUPITER is, “a bit much even for a large state like North Rhine-Westphalia,” in reference to the substantial costs involved. In his speech, he emphasized North Rhine-Westphalia’s consistent path from coal to AI, pointing out that the inauguration of JUPITER also marks a milestone for structural transformation in the Rhineland region. Finally, Wüst picked up on the Chancellor’s train of thought, and his closing words sound like a solemn commitment: “JUPITER will have its home in Jülich, here with us in North Rhine-Westphalia, here with us in Germany, but it is a gain for the whole of Europe,” he said.

JUPITER RISING – the moment of the "initial spark"

The big moment of the ceremonial “startup” took place in the middle of the program. Between the first and second panels, host Döbbelt took the floor and announced simply that it is now time “to officially inaugurate JUPITER once again.” On a day like this, for a system like this, there are no red and white ribbons to cut or the usual groundbreaking ceremony – a “buzzer” is set to make history. As the countdown reaches zero and the button is pressed, fanfare sounds, and on a large video screen, JUPITER begins to come to life: numbers race across the displays, light floods the Modular Data Centre (MDC) that houses the exascale supercomputer, and then it’s official: Jülich “switched on” JUPITER. A musical crescendo accompanied the moment – lights, camera, action: a new star is born in Europe’s research sky.

The system is about far more than computing power: It is about technological sovereignty and competitiveness, about AI advancements in Germany, about excellent research, and about, “playing a crucial role in shaping this technological revolution, as a nation both with a long scientific tradition and as an industrial nation,” as Merz put it.

What the guests take away

From the many contributions and conversations, the inauguration laid bare that JUPITER stands for much more than technological progress and energy efficiency. From the European perspective, it provides crucial research support to researchers whose work contributes to solving global challenges and advancing our understanding of life, the environment, and the universe.

At the same time, it is evident that JUPITER is just the beginning. Thanks to its modular design, it will be continuously developed – with components such as JARVIS, an inference module that will enable the “JUPITER AI Factory” and create a direct link to industry. The “Cluster” module, which will be added soon, is specifically designed for training artificial intelligence. With JSC being able to play the role of a cornerstone in Germany’s application for a European AI GigaFactory, Jülich is already actively seeking ways to further leverage its new flagship system in this rapid expansion stage of European computing infrastructure. *Lisa Maiburg*

To read the original press release, please visit: https://www.fz-juelich.de/en/news/archive/press-release/2025/ceremony-jupiter

The inauguration was held in German. Quotations from the event have been translated from the German original.

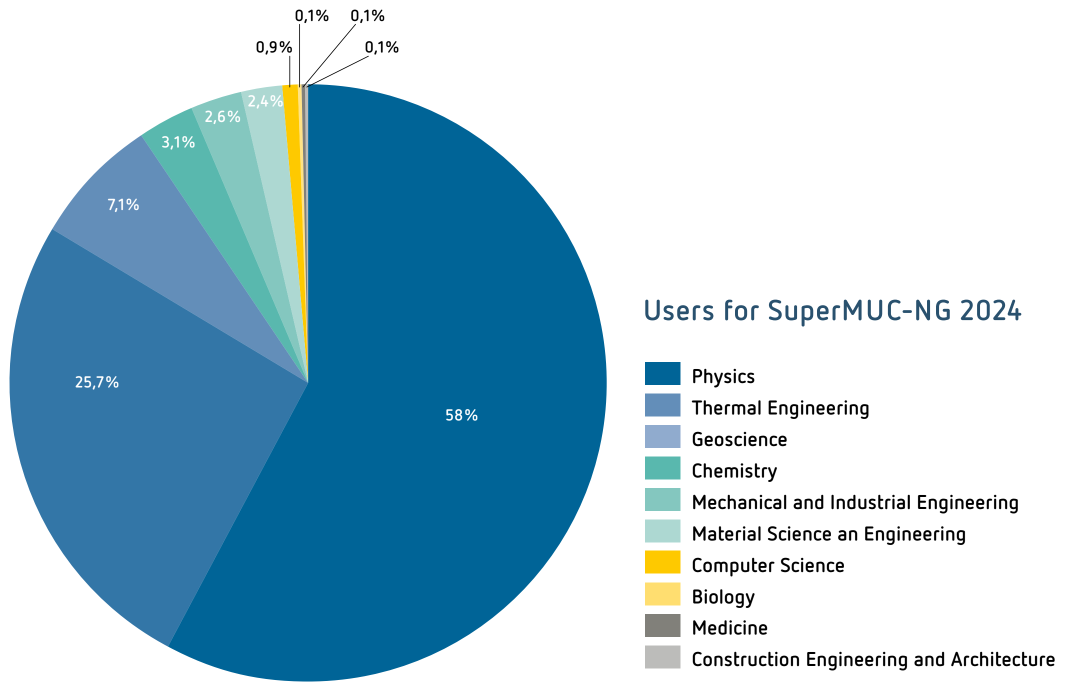
25 years of HPC at LRZ: Technical and Scientific Highlights Across a Quarter Century of Computing

For 25 years, the center has been a major part of high-performance computing at the national level. Time for a walk through LRZ's HPC history.

The HLRB II system at LRZ. Inaugurated in 2006, it was an SGI Altix 4700 and the first system to be deployed at the newly built data centre on the campus in Garching.

In 1990, the United States-based market research firm Superperformance Computing Services conducted a survey of leading computer manufacturers. They subsequently predicted that by the turn of the millennium, “monster computers” with more than 256 processors, each no larger than 16 square centimeters, would be capable of performing trillions of floating-point calculations per second. The reality was different. According to the Top500 list, the most powerful supercomputers in 2000 achieved 2,000 to 3,000 GFlops per second (Gflops/s), or about two to three trillion operations.

Among the world’s top supercomputers, a German system ranked fifth with a peak performance of around 2,016 GFlop/s. The HLRB I – the acronym stands for “High Performance Computer Bavaria” – was powered by 1,344 SR8000-F1 processors from Hitachi. Although the HLRB I was not the first supercomputer commissioned by the Leibniz Supercomputing Centre (LRZ), it was the first to propel the Bavarian computing center to national HPC center status in Germany. This was a milestone in LRZ’s history, as it became more European and international. The LRZ now boasts a number of technical highlights that



Statistics of SuperMUC-NG usage for 2024 demonstrate the strikingly diverse user community for LRZ high-performance systems. The institute has long taken pride in the diverse user communities served by all of its national machines.

SuperMUC Phase 1 and 2 ran from 2012 to 2018. Phase 1 started out as Europe's fastest system and was the first hot-water cooled supercomputer worldwide when entering the Top500 in 2012.

have since become standard in supercomputing. For instance, LRZ developed water cooling technology with IBM and later optimized it continuously with Lenovo. The center's current flagship system, SuperMUC-NG Phase 2, is the first system that cools exclusively with hot water. In collaboration with IQM Quantum Computers, LRZ integrated a quantum system into the current supercomputer – a significant step forward in HPC technology. In fact, all four HPC clusters built in Bavaria – the HLRB systems followed by the SuperMUC series of systems – were among the world's top 10 most powerful supercomputers when they went into operation.

A-ha! moments for research

While biannual Top500 lists represent snapshots of computing power through the years, these lists also show that the highest rankings are quickly lost because the global scientific community's demand for computing power and storage capacity is constantly growing. At LRZ, computing performance has approximately doubled every 14 months, while storage space has doubled every half year. Similar developments show up and down the Top500 list.

Supercomputers attract less sustained interest than the people who use them. All four LRZ systems have advanced world-class research and enabled "a-ha!" moments among scientists. SuperMUC-NG simulated blood flow in veins and the brain for the European Centre of Excellence, CompBioMed, and searched libraries for candidate compounds to use for fighting the COVID-19 pandemic. For the Technical University of Munich (TUM), SuperMUC-NG calculated a model of the human lung that changed our understanding of the organ. Since then, the lung has been regarded not as a structure made up of grape-shaped sacs but as a sponge-like structure that, when unfolded, is larger than the skin.

In 2013, the 1Kite project used SuperMUC to organize genomic data and reconstruct the family trees of insects and birds. These biological fundamentals are still valid and are among the most frequently cited results in the field. For the international astrophysics project Millennium TNG, LRZ's HPC systems simulated dark matter multiple times. This led to the discovery and measurement of neutrinos, among other advancements in the astrophysics and cosmology community.

Further usage data from the past 25 years shows that simulation and modelling have become the third pillar of the natural sciences, alongside experiments and theory. At LRZ, which boasts a variety of different computing architectures, supercomputers have never only been used exclusively by traditional heavy users in fields like

The center's current flagship system, SuperMUC-NG Phase 2, is the first system that cools exclusively with hot water.

astrophysics, thermodynamics, chemistry, and engineering, but biologists, pharmacologists, climate researchers, mathematicians, computer scientists and historians use them. Artificial intelligence methods are already further expanding this community, and will continue to in the coming years, allowing the number of interdisciplinary projects to grow.

National collaborations lead into the next era of supercomputing

During the past 25 years of HPC at LRZ, the center has partnered with Germany's other leading centers, establishing the Gauss Center for Supercomputing (GCS) in 2007. Since then, the three national computing centers in Germany have pooled their HPC resources. During this time, Dennard scaling ended, and Moore's Law reached its limits. Consequently, it is no longer possible to increase computing power and storage space simply by creating denser transistors or by parallelizing or by parallelizing more processors, especially since supercomputers' energy requirements are gradually exceeding data centers' capabilities.

As is the case everywhere else, the future of supercomputing at LRZ will consist of experimentation and research. The expansion of the current high-performance computer, SuperMUC-NG, includes nearly 27,000 CPUs and around 1,000 GPUs, which accelerate computing and enrich HPC with statistical methods. The system is also connected to quantum computers, opening up new possibilities. In July 2025 the LRZ installed its first photonic processors. These processors use light instead of electrical pulses for computing and are currently undergoing testing to determine their compatibility in an analog-digital architecture for supercomputers.

Not every prediction about the future is reliable. However, in addition to innovative hardware, the LRZ will rely on new programming models, software stacks, and the ideas of building management specialists to improve the performance and, more importantly, the energy efficiency of the next supercomputers. *Susanne Vieser*

EU Commissioner Visits HLRS and JSC on Tour of AI Factories

Henna Virkkunen, the European Commission Executive Vice President for Tech Sovereignty, Security, and Democracy, recently visited the High-Performance Computing Center Stuttgart and Jülich Supercomputing Centre to get a closer view on the centers' roles in bolstering made-in-Europe artificial intelligence innovation.

In the last year, all three Gauss Centre for Supercomputing (GCS) centers were selected to participate in one of the EuroHPC Joint Undertaking's (EuroHPC JU's) 19 AI Factories. The initiative brings together high-performance computing (HPC), data infrastructure, and expertise in artificial intelligence (AI) to support increased uptake of AI for academic and commercial researchers in Europe. The High-Performance Computing Center Stuttgart (HLRS) and the Jülich Supercomputing Centre are both coordinating AI factories – HammerHAI in Stuttgart and JAIF in Jülich – with the Leibniz Supercomputing Centre (LRZ) serving as one of the partners in HammerHAI.

HammerHAI is focused on providing AI-optimized computing infrastructure and solutions for research and industry, with a particular focus on automotive, engineering, manufacturing, and global challenges research. HammerHAI is also working closely with start-up companies and small and medium-sized enterprises to ensure that the benefits of generative AI are accessible to Europe's large corporations and small businesses alike. This includes offering personalized, expert support for AI users at all stages in the AI lifecycle. The three-year initiative is funded with approximately €85 million.

Beginning in 2026, HLRS will also host a new AI-optimized EuroHPC JU supercomputer in Stuttgart on behalf of HammerHAI that can run large-scale AI applications. A procurement for the new system is currently underway, the results of which are expected to be announced later this year.

"HammerHAI represents an important next step in the evolution of HLRS's strategy, as we have a long history of collaboration not only with science but also with industry," said Dr. Bastian Koller, Managing Director of HLRS and lead coordinator of the HammerHAI consortium. "Since launching HammerHAI, we have received many inquiries and look forward to the success stories that this AI factory will help to produce."

The JUPITER AI Factory (JAIF) at the Jülich Supercomputing Centre also brings together AI expertise, advanced computing, data, and AI infrastructure, and academic and commercial researchers, to develop AI solutions for research and industry in strategic key sectors, including healthcare, energy, climate change and environmental science, education, culture and media, the public sector, finance and insurance, and manufacturing. It will enable efficient use of Europe's first exascale supercomputer, JUPITER, and provide cloud-based inference resources via JUPITER's inference model, JARVIS. The project runs for three years and is funded with €55 million by the EuroHPC Joint Undertaking, the German Federal Ministry of Research, Technology and Space, and the Ministries of Science of North Rhine-Westphalia and Hesse.

EU Commissioner's visits highlight the importance of AI Factories, made-in-Europe technology

In July, HLRS welcomed Henna Virkkunen, the European Commission Executive Vice President of Tech Sovereignty,

"Since launching HammerHAI, we have received many inquiries and look forward to the success stories that this AI factory will help to produce."

Henna Virkkunen, European Commission Executive Vice-President for Tech Sovereignty, Security and Democracy

Security, and Democracy, on her tour of all AI factories. "AI Factories like HammerHAI will be critical in fostering a German and European ecosystem for generative AI and a key asset for developing European competitiveness in AI and strengthening Europe's digital sovereignty," she said in conjunction with the event.

Already in April, Virkkunen had visited JSC as the first stop on her tour through Europe, emphasizing the role that the JAIF ecosystem would play in supporting AI development in Europe.

"This cutting-edge facility, to be inaugurated in just a few months, is more than just a milestone; together with other AI Factories it is the cornerstone of our AI continent strategy to foster German and European AI innovation and create a step-change in EU competitiveness," she said. "As I embark on my tour of Europe's emerging AI factories, starting right here in Jülich, I am inspired by its creativity and transformative potential. The future of Europe as AI Continent is being forged right here, right now."

Virkkunen's visits also highlighted how the European Union (EU) plans to expand the AI factory models to build so-called "Gigafactories" for AI. Seventy-six institutions applied to set up AI Gigafactories across the EU this year, with award announcements coming at the end of 2025. The AI Gigafactories will expand the AI Factory model by providing extreme-scale computing resources to awardees that will further accelerate AI model development, training, and deployment. *Eric Gedenk*

For more information on HammerHAI, please visit:
https://www.hammerhai.eu/

For more information on JAIF, please visit:
https://www.fz-juelich.de/en/jsc/jupiter/jaif-jupiter-ai-factory

For more information on the AI Factory initiative, please visit:
https://digital-strategy.ec.europa.eu/en/policies/ai-factories

As coordinator of HammerHAI, HLRS will host a new AI-optimized supercomputer for the EuroHPC Joint Undertaking beginning in 2026. HLRS's Hunter supercomputer has already begun supporting AI applications. L-R: Michael Resch, Henna Virkkunen, and University of Stuttgart Rector Peter Middendorf.

After a certain point in the development of a turbulent flow — for example, as air moves over a wing in flight — the outer region of the turbulent boundary layer maintains a persistent, self-similar physical structure.

SCIENCE
FEATURES

Large-Scale Simulation Reveals Novel Insights about Turbulence

Using HLRS's Hawk supercomputer, University of Stuttgart scientists have for the first time produced high-resolution simulation data characterizing the transition from low to high Reynolds numbers in turbulent boundary layers.

Scientists at the University of Stuttgart's Institute of Aerodynamics and Gas Dynamics (IAG) have produced a novel dataset that will improve the development of turbulence models. With the help of the Hawk supercomputer at the High-Performance Computing Center Stuttgart (HLRS), investigators in the laboratory of Dr. Christoph Wenzel conducted a large-scale direct numerical simulation of a spatially evolving turbulent boundary layer. Using more than 100 million CPU hours on Hawk, the simulation is unique in that it captures the onset of a canonical, fully-developed turbulent state in a single computational domain. The study also identified with unprecedented clarity an inflection point at which the outer region of the turbulent boundary layer begins to maintain a self-similar structure as it moves toward high Reynolds numbers. The results appear in a new paper published in the *Journal of Fluid Mechanics*.

"Our team's goal is to understand unexplored parameter regimes in turbulent boundary layers," said Jason Appelbaum, a PhD candidate in the Wenzel Lab and leader of this research. "By running a large-scale simulation that fully resolves the entire development of turbulence from an early to an evolved state, we have generated the first reliable, full-resolution dataset for investigating how high-Reynolds-number effects emerge."

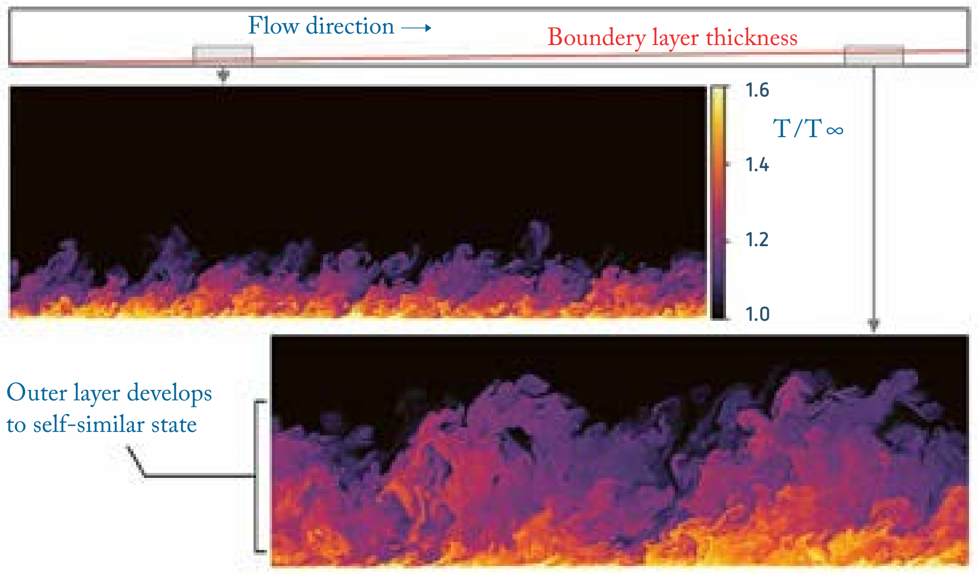
Why it's difficult to study moderate Reynolds numbers

During flight, turbulence causes very high shear stress on the surface of an aircraft. The resulting drag can reduce flight performance and fuel efficiency. To predict this effect, aerospace engineers rely on computational models of the turbulent boundary layer, the millimeters-thin region where the surface of the aircraft interacts with free-flowing air.

> "Our team's goal is to understand unexplored parameter regimes in turbulent boundary layers."
> Jason Appelbaum, PhD candidate in the Wenzel Lab

For industrial applications, turbulence models do not need to replicate physics down to the finest details; they must only be accurate enough for practical use and be capable of running smoothly on modest computing resources. Before engineers can use such simplified models, however, scientific research using high-performance computing systems is necessary to provide the data on which they are based. This is why the Wenzel Lab has long used HLRS's supercomputers to run NS3D, direct numerical simulation software it created to investigate fundamental physical properties of turbulent boundary layers at extremely high resolution.

Scientists in the field of computational fluid dynamics (CFD) use a figure called the Reynolds number to characterize the developmental state of a turbulent boundary layer. The Reynolds number is the ratio of inertial forces to viscous forces in a fluid flow, which governs the local range of turbulent eddy sizes. At low Reynolds numbers, which occur early in the motion of a surface through air, nonlinear convective instabilities responsible for turbulence are quickly damped by viscous action at small scales. With increasing Reynolds number, the turbulent boundary layer



The top of this figure represents the IAG teams large-scale simulation, which captured the complete development of a turbulent boundary layer from low to high Reynolds numbers. As a flow moves across a surface, the outer region of the turbulent boundary layer becomes thicker. Beyond a certain point, it retains similar physical properties.

becomes thicker. Large, coherent structures emerge, creating a more complex turbulent system that is not simply an extrapolation of trends at low Reynolds numbers, but has its own distinct properties.

In the past, CFD simulations have generated rich datasets for understanding turbulence at low Reynolds numbers. This is because the computational domain size and the necessary number of simulation time steps involved at this stage are still relatively small. By today's standards, this means that the simulations are not prohibitively expensive. Laboratory experiments also provide invaluable data for turbulence research. For quantities relevant to the present study, however, they have only focused on the high Reynolds number regime due to physical limitations. Sensors can only be machined so small, and some fundamental physical quantities, such as shear stress, are notoriously difficult to measure in the lab with high accuracy.

As a result, scientists have accumulated a wealth of simulation data for low Reynolds numbers and reliable experimental data for high Reynolds numbers. What has been missing, however, is a clear picture of what happens in between, as both simulation and experimental methods are of limited use. Appelbaum and his collaborators in the Wenzel Lab set out to attack this problem directly.

A sharp turn

Using HLRS's Hawk supercomputer, Appelbaum ran a series of simulations that, when the results were stitched together, replicate the entire evolution of a turbulent boundary layer from low to high Reynolds numbers.

Although the "real-life" situation the simulation represented might seem vanishingly small – traveling at Mach 2 for approximately 20 milliseconds – the campaign required large-scale computing power. The team used 1,024 computing nodes (more than 130,000 cores) on Hawk – one-fourth of the entire machine – for hundreds of short runs, each of which lasted 4 to 5 hours. In total, the simulation required more than 30 days of computer runtime.

"Most research groups would not take the risk of spending so much computational time on a problem like this, and might instead look at other interesting research problems that aren't as expensive," Appelbaum said. "We're the weird ones who put all of our eggs in this one basket to investigate a long-standing gap in the research."

The investment paid off. In their large-scale simulation the investigators focused on (among other factors) the skin friction coefficient, a value that represents the proportion between shear stress at a solid surface in a moving fluid in comparison to free momentum of the flow. It is a key parameter describing the shape of the mean velocity profile and is fundamental in determining the viscous drag.

Appelbaum used the results of the simulation to show how the previously separate datasets for low and high Reynolds numbers blend together. Whereas past research could only estimate through interpolation how the datasets might intersect, the IAG team's results reveal a sharp turn. Notably, they identified a change in skin friction scaling that is linked to the establishment of a fully-developed state in the outer 90 percent of the boundary layer. This self-similar state is a milestone in the turbulent boundary layer's development, signaling that scaling behavior continues in a predictable way as it evolves to industrially relevant Reynolds numbers.

"To understand self-similarity, it helps to think about the aspect ratio of a photograph," Appelbaum explained. "If I have a rectangular picture where the lengths of the sides have a ratio of 1:2, it doesn't matter whether the picture is the size of my hand or if I scale it to the size of a bus. The relationships among the elements in the photo remain self-similar, regardless how large it is. Our work confirms that the outer region of the turbulent boundary layer takes on the same kind of self-similarity once the system reaches a specific Reynolds number. Importantly, this state is coupled with the change in scaling behavior of the skin friction coefficient, which experiments have shown to remain in effect until very high Reynolds numbers seen in aerospace applications. This allows us to get an early, but realistic glimpse of the turbulent behavior in this ultimate regime of turbulence."

Increased performance offers new opportunities for research and engineering

This new dataset offers a unique resource that will better enable researchers in the computational fluid dynamics community to investigate turbulent boundary layers at moderate Reynolds numbers. For the Wenzel Lab, the next step will be to dive deeper into the physics behind the inflection point they identified. Appelbaum says that the team already has some ideas about this and plans to publish a follow-up paper soon.

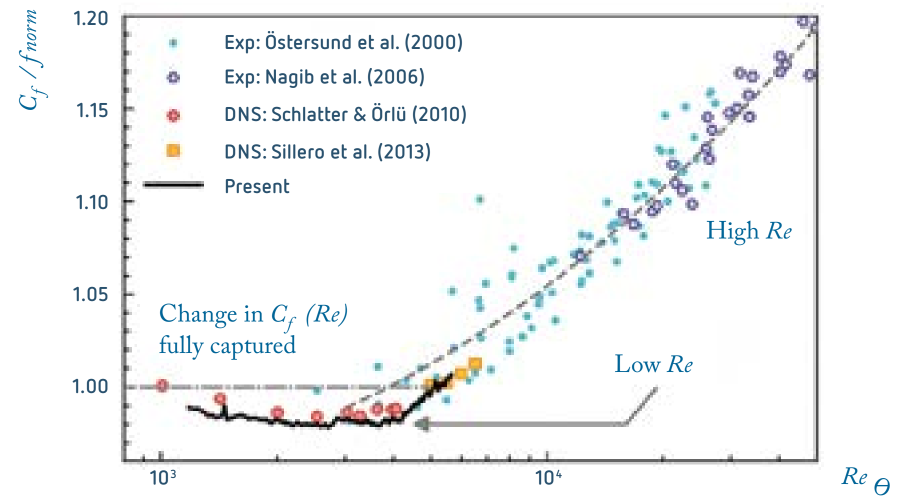
In other ongoing work in the Wenzel Lab, the scientists have been busy porting the NS3D code to GPUs on HLRS's newest supercomputer, Hunter. With the help of user support staff at HLRS and computing processor manufacturer AMD, they have already verified that the code remains physically accurate and performant using the new, GPU-accelerated system. In the coming months they will be optimizing NS3D to ensure that it takes full advantage of the increased performance that Hunter offers.

"We anticipate being able to to simulate larger domains at even higher turbulent states," Appelbaum said. "More computing performance will also make it more feasible to do studies in which we might run several simulations to investigate the scaling behavior of two or more parameters simultaneously."

In work that points toward this future, Tobias Gibis, a member of the Wenzel Lab and co-author of the present work, recently defended his thesis, in which he unified the scaling behavior of heat transfer and pressure gradients in turbulent boundary layers. Appelbaum added, "Building on Christoph and Tobias's transformative work on the influence of heat transfer and pressure gradients to include Reynolds number effects would undoubtedly have very high scientific value. The support and resources from HLRS are the bedrock for this type of heavy computational work."

In the meantime, the team's dataset at moderate Reynolds numbers will contribute to a pool of wall-bounded turbulent flow data and could aid in the development of more comprehensive, more accurate turbulence models. This will give engineers new capabilities for optimizing aircraft designs for a wider range of operating conditions, and for improving other kinds of machines like fans or automobiles whose efficiency relies on managing effects in turbulent boundary layers. *Christopher Williams*

Related publication: Appelbaum J, Gibis T, Pirozzoli S, Wenzel C. 2025. The onset of outer-layer self-similarity in turbulent boundary layers. *J Fluid Mech.* 1015: A37.



While past research produced data reflecting the skin fraction coefficient (cf) at low and high Reynolds numbers, a large-scale simulation on HLRS's Hawk supercomputer for the first time revealed a specific period at moderate Reynolds numbers at which a fundamental change occurs.

JSC Researcher Improves Deep Learning Training

University of Iceland PhD student Marcel Aach is a research assistant at the Jülich Supercomputing Centre. As part of his graduate research, he used JSC computing resources to develop a new method for optimizing the hyper-parameters needed to effectively train new artificial intelligence models.

Within the expanding field of artificial intelligence (AI), computer scientists continuously develop new AI models that boost scientific, commercial, and personal productivity. The need for more tailored models is rapidly growing, so scientists must optimize training these models that are increasingly demanding more computational resources.

Specifically, the training of large, complex neural networks used for deep learning tasks like advanced image classification or drug design is becoming prohibitively expensive, which means researchers are reliant on the world's most powerful computational resources to continue to develop new, powerful models.

Marcel Aach has focused his academic and professional experiences on one aspect of AI training that has big implications for improving the efficiency of training large models for deep learning tasks – optimizing the so-called hyperparameters that control the accuracy and speed of models and their training.

In his recently published doctoral thesis, Aach used Jülich Supercomputing Centre's (JSC's) JUWELS Booster supercomputer to develop a new hyperparameter optimization (HPO) method to train the next generation of AI models more efficiently, specifically by progressively shifting computational resources toward the most important training tasks. "For a long time, researchers have known that if you want to improve an AI models performance, you have to increase the amount of training data you have," he said. "At this point, though, it means that training state-of-the-art AI models requires massive amounts of training data and needs intensive computational resources. Working on national and European research projects at JSC, I realized that HPO is a topic that basically unites everyone who builds machine learning models, no matter the research domain." The key advantage of HPO is that, when applied effectively during training, it significantly improves efficiency and enhances the quality of resulting AI models without requiring additional data or computational resources.

> "For a long time, researchers have known that if you want to improve an AI models performance, you have to increase the amount of training data you have."
> Marcel Aach, PhD student, University of Iceland

Hyperparameter optimization speeds up training with help of GPUs

Aach first studied AI through his experiences optimizing high-performance computing (HPC) simulations to run in parallel. With this method, the calculations of a specific simulation are distributed across many computer cores, and the results are shared efficiently across a supercomputer. As Aach became more interested in neural networks, he applied his HPC experience to AI training and realized that exploiting parallelism when training a deep learning model was essential to improve efficiency.

He likened hyperparameters to knobs and levers that scientists can set before training a model. They determine the speed, specificity, and level of detail, among other aspects, that researchers use in neural network training. "For these large models, it is essential to pick good hyper-parameters from the beginning, because you do not want more of these costly GPU trainings than absolutely necessary," he said. "We want to make sure we're helping build the best models, but hopefully with one or a couple of training runs rather than dozens."

With his access to JUWELS Booster and experience optimizing large-scale computational codes, Aach knew he could use the system's 3,744 GPUs to increase parallelism in two ways: by testing different hyperparameters for a model and by completing test runs as efficiently as possible.

He noted that no matter what the training goal is, all scientists who are training for deep learning tasks must evaluate a model's performance on a variety of scales and with various hyperparameters. If these training runs are running in parallel, scientists can more quickly determine the best performing hyperparameters for the main model training. To more efficiently complete these test runs, researchers must also ensure that their models are optimized to run across many more GPUs than in the past.

While these two areas of parallelism already improve efficiency, Aach wanted to develop a technique that could further improve HPO by shifting computational resources to focus on the most-promising models. "For my doctoral research, I developed a way for us to study the various test runs of a neural network, focus on each iteration's accuracy or error rate, and then shift GPUs from the networks that were not performing well to the models that were performing the best," he said. "So we mixed these two ways of parallelism and ultimately got results in a faster way by consolidating resources for the best models."

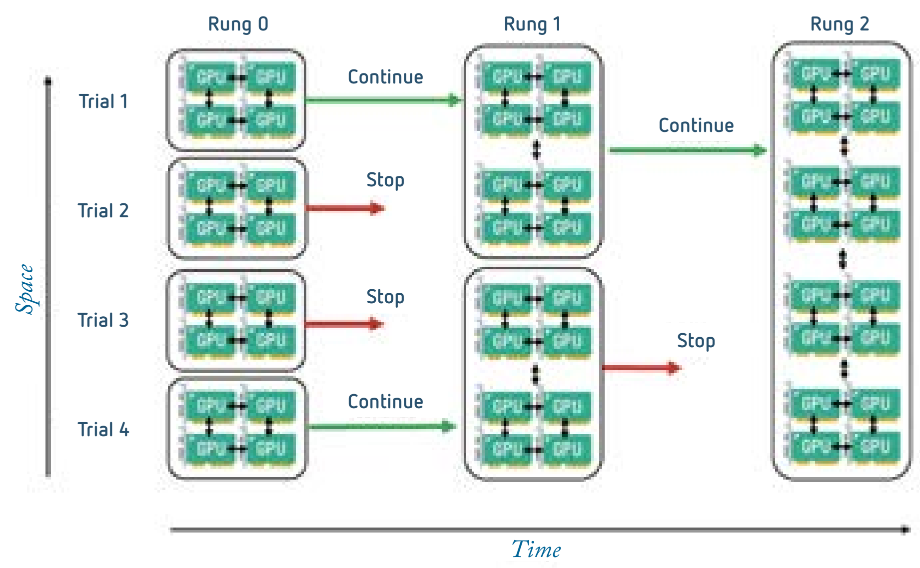
Aach tested his method on a variety of deep learning applications, ranging from computational fluid dynamics and remote sensing to computer vision benchmarks and data analysis tasks using the OpenML machine learning platform. Results showed that his method improved training efficiency across all applications.

JSC's exascale resources support larger-scale model training

With his PhD successfully defended, Aach is next shifting into a postdoctoral research position at JSC to continue focusing on HPO. His work comes at an opportune time, as JSC just inaugurated JUPITER, its next generation supercomputer that will be Europe's first exascale system. JUPITER not only represents a more than ten-fold increase in computing power over JUWELS, but also has roughly 24,000 next-generation NVIDIA Hopper GPUs – offering a massive performance boost for AI model training.

"We evaluated this method on petascale computing systems like JUWELS, using roughly 1,000 GPUs, several terabytes of training data, and a few million training parameters," Aach said. "With the arrival of JUPITER, we can scale that up to a few hundred terabytes of training data and billions of parameters being trained for a given neural network. I'm looking forward to working at JSC and seeing how my method performs when we bring it to this scale." *Eric Gedenk*

You can read Aach's full thesis at: https://opinvisindi.is/handle/20.500.11815/5293



Marcel Aach focused his PhD thesis on how to improve AI training. The chart illustrates the crux of his work: Aach designed a way to shift GPU resources toward the most promising training instances.

SCIENCE
FEATURES

Researchers Simulate Astrophysical Turbulence at Unprecedented Detail

An international research collaboration recently used the SuperMUC-NG supercomputer to run the largest, most detailed simulations of magnetism and turbulence in the interstellar medium.

High-performance computing (HPC) has become an indispensable tool for astrophysics. Researchers who want to better understand our universe often turn to supercomputing to model phenomena in our galaxy either too large or too complex for us to probe otherwise.

Recently, a research team led by James Beattie, a postdoctoral researcher at Princeton University and the Canadian Institute for Theoretical Astrophysics at the University of Toronto, focused on better understanding how magnetic fields influence the highly turbulent motions through the interstellar medium (ISM). Using HPC resources at the Leibniz Supercomputing Centre (LRZ), the team was able to model turbulence in the ISM in unprecedented detail, calling long-held assumptions of the role of magnetic turbulence into question in the process and providing new research directions for next-generation experiments in space. The team published its work in Needs to be italicized.

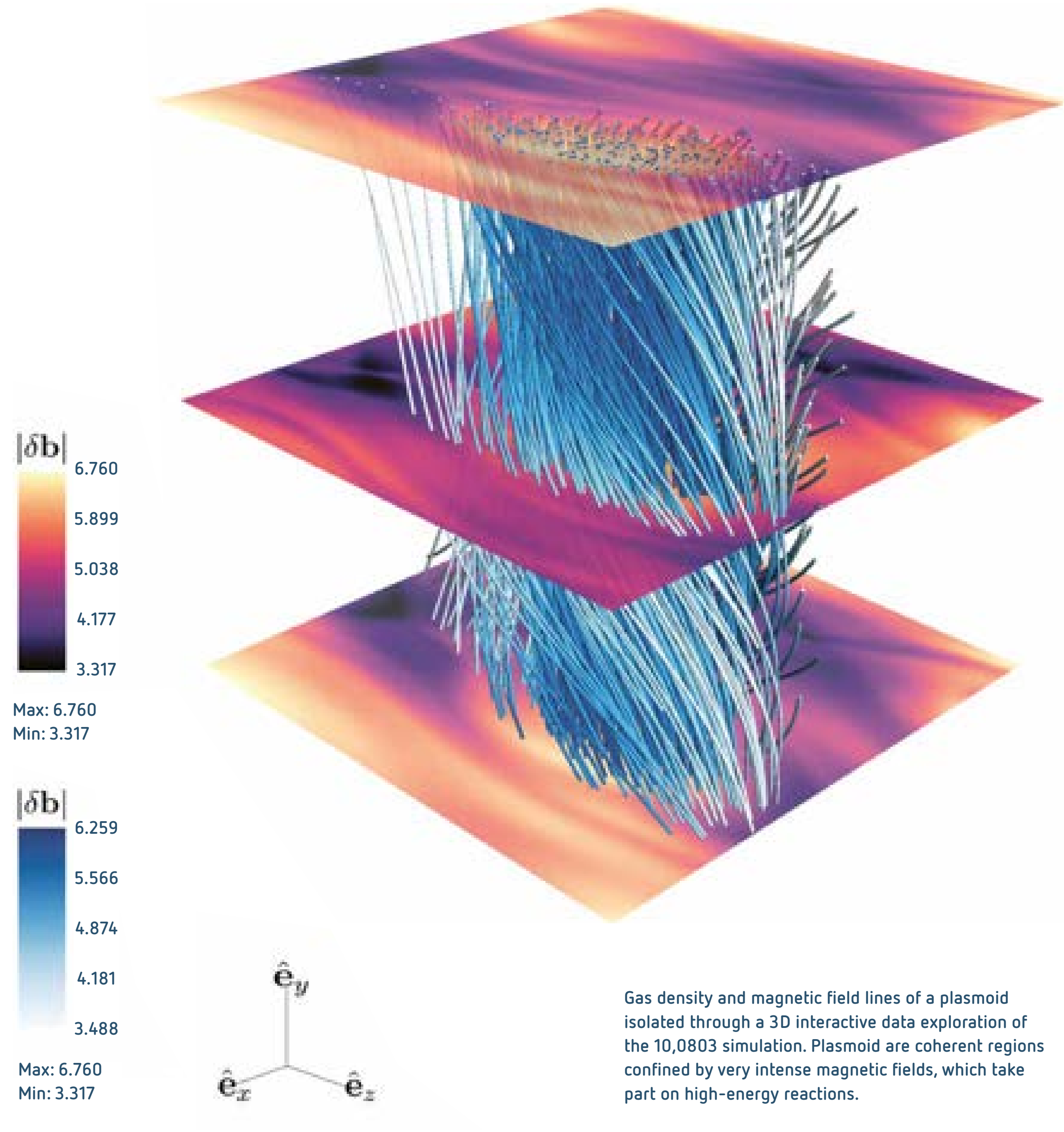
"The ISM is ultimately the source from which new stars are formed," said Salvatore Cielo, user support specialist in LRZ's Astrolab and co-author on the paper. "Understanding how this process works exactly is a longstanding problem in astrophysics, one where theory alone cannot yet explain from first principles. In fact, our simulations managed to predict a behavior of the magnetic energy that has evaded other theoretical models."

In addition to Princeton University, the University of Toronto, and LRZ, the team received support from researchers at the Australian National University and Heidelberg University. LRZ is one of the three HPC centers that comprise the Gauss Centre for Supercomputing (GCS).

Capturing the many scales of turbulence simulations

Physicists consider turbulence one of the last major unsolved challenges in their discipline, and for good reason – turbulent motions of fluids are inherently chaotic and trying to accurately study or simulate how these chaotic motions influence one another at all scales has proven elusive. Moreover, the turbulent vortices fragment into ever smaller ones, with a subsequent "cascade" from large to small structures.

> **"The ISM is ultimately the source from which new stars are formed."**
> Salvatore Cielo, LRZ



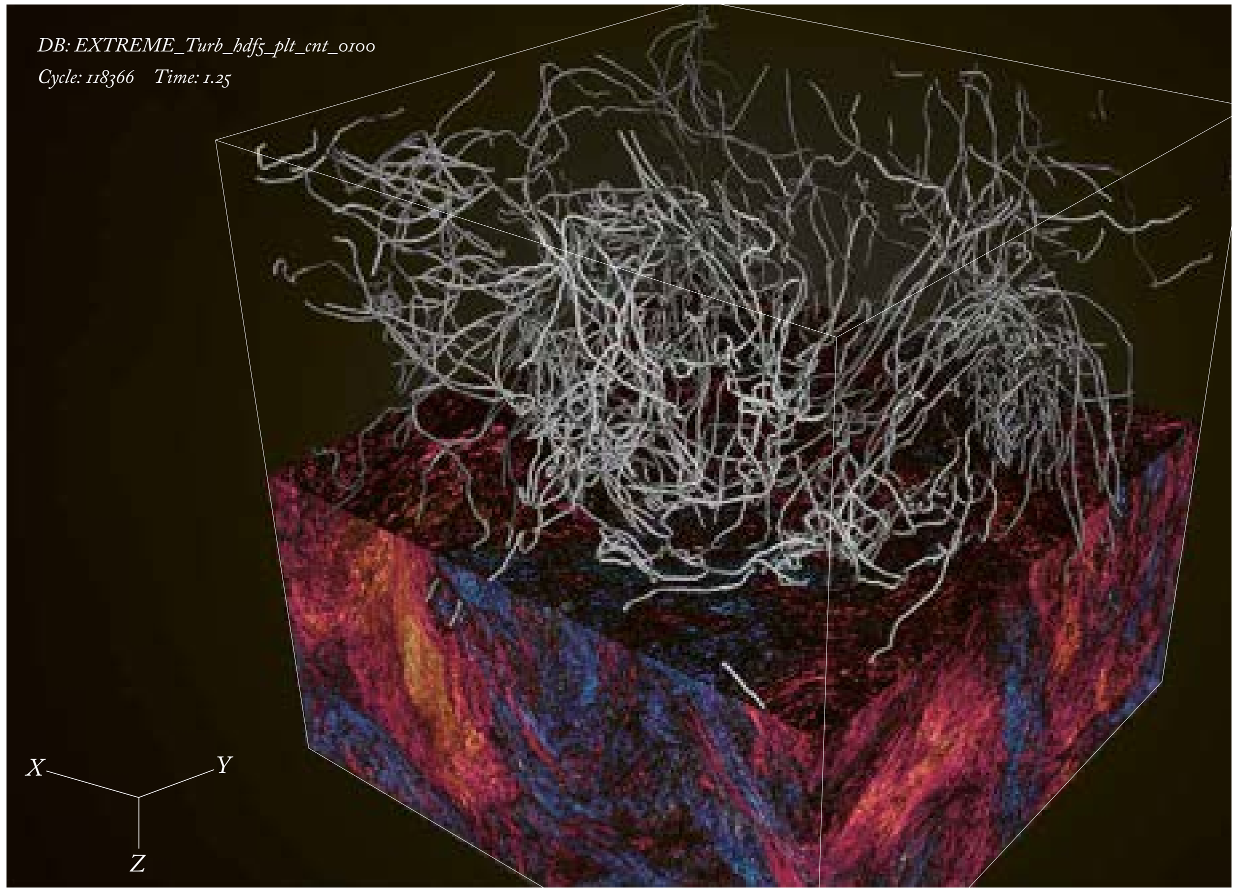
Gas density and magnetic field lines of a plasmoid isolated through a 3D interactive data exploration of the 10,0803 simulation. Plasmoid are coherent regions confined by very intense magnetic fields, which take part on high-energy reactions.

The ISM adds even more challenge: turbulence happening in the galaxy is also influenced by magnetic fields. While magnetic fields from far-away stars may be weak, turbulent vortices in the ISM amplify them like a spinning dynamo. To model such a pervasive, fluid-like structure like the ISM, researchers use magnetohydrodyanmics (MHD) simulations. On a computer, the researchers create a grid, then solve the MHD equations within each grid space, which tell them how magnetic fields, particles and gas in the ISM influence one another. The finer the grid's mesh, the more physics they can fold into the simulation. But more grid points mean higher computational costs. "Turbulence in the ISM does not have any defined geometry or symmetry that you can use to simplify the modeling, so you really have to optimize your code and tackle these calculations with a brute-force computational approach," Cielo said.

To get a more accurate view of how magnetic turbulence influences the ISM, the team created a high-resolution grid that consisted of 10,080 units per dimension – an ISM portion about 30 light-years per side, large enough to encompass structure such as extreme dense regions and voids, plasmoids (a form of confined magnetic "bubbles"), and to provide statistics on how the magnetic fields align with the flow. But the model can also be scaled down to study finer phenomena, such as wind and plasma ejections coming from the sun, which influence space travel or satellites.

As a member of LRZ's Astrolab, Cielo works closely with researchers who won allocations of computing time for astrophysics research. Working with Beattie's team, Cielo supported the researcher by helping to optimize the code to run across 140,000 cores on LRZ's SuperMUC-NG Phase 1 system.

The team used mixed-precision algorithms – an optimization used also in machine learning and artificial intelligence applications – to reduce the load in memory in less-critical spaces in the simulation, allowing the cores to direct more computational resources to expanding the simulation size. Cielo also led an interactive, three-dimensional visualization of the data, which also required hundreds of nodes on SuperMUC-NG Phase 1.
This visualization, along with a detailed 2D visualization created by Beattie, gave the team new insights into how magnetic fields alter energy cascading through the ISM, suppressing some of the smallest-scale turbulent motions. In turn, they enhance Alfvén waves, or low-frequency disturbances powered by the magnetic fields in the cosmos. These findings show the researchers that magnetic fields may play a bigger role in stabilizing the ISM against excessive fragmentation that previous models have suggested and gives scientists a new line of inquiry into more accurate magnetic turbulence simulations in the future.
"With this extremely high-resolution work, we are able to make qualitatively new predictions that take us another step closer to unraveling the mystery of stability and energy transport in the ISM," Cielo said. "The problem of magnetized turbulence touches fundamental chords in several branches of physics, so we hope that many other scientists will find this work inspiring. We encourage curious people to read the paper, read about the research happening on our groups' sites, and contact us about exciting new research projects for SuperMUC-NG."



Interactive 3D visualization on SuperMUC-NG phase of the full simulation box, at a resolution of 5,0403, showcasing the different scales of turbulent motions. Bottom half: volumetric rendering of the gas density of the full turbulent box; top half: streamlines of the magnetic field.

"SuperMUC-NG Phase 1 and 2 are well-suited for these kinds of simulations, and with a great research team and a good code, you can scale up to new computational heights."
Salvatore Cielo, LRZ

GPU accelerators promise more efficient modeling

Cielo continues to support the team as it uses its findings to inform new, exciting simulations running on LRZ resources. As the center finishes acceptance testing of SuperMUC-NG Phase 2, a significant upgrade to the center's flagship system, the team looks forward to porting its application to take advantage of GPUs available on the new machine. "This is another aspect of supporting our users for the next generation of HPC – it requires large efforts to port software to run on GPUs; there is no free lunch," Cielo said.

Cielo and the research team have already begun exploring several alternatives for porting this and other applications to take advantage of GPU acceleration or selecting other GPU-ready ones. The team is also in the early stages of exploring how to use AI-accelerated algorithms that would help free up precious memory during a simulation primarily limited by information coming on and off memory.

For Cielo, the team's focused work has created a strong foundation from which to work, and he and other support staff in LRZ's Astrolab will continue to provide optimization and visualization support to only sharpen the team's insights gleaned from its recent work. "They run very optimized software, and we help create the perfect computational environment for the work – the software, compilers, and everything needed to run well on our hardware," Cielo said. "SuperMUC-NG Phase 1 and 2 are well-suited for these kinds of simulations, and with a great research team and a good code, you can scale up to new computational heights." *Eric Gedenk*

Related Publication: Beattie, J.R., Federrath, C., Klessen, R.S. et al. "The spectrum of magnetized turbulence in the interstellar medium." *Nat Astron* (2025). https://doi.org/10.1038/s41550-025-02551-5

The Grace Hopper superchip is the basis for the current test system and will prepare LRZ and its users for the upcoming Vera Rubin architecture.

Warming Up for Blue Lion

As part of its collaborative efforts in advance of its Blue Lion system, LRZ and vendor HPE installed a testing and optimization platform to get users and staff alike prepared for a new, powerful architecture.

Blue Lion sends in the young ones: For testing and training purposes, Hewlett Packard Enterprise (HPE) has installed Blue Cubs at the Leibniz Supercomputing Centre (LRZ), a first test installation to prepare for the launch of Blue Lion, LRZ's next supercomputer. This will allow experimentation, initial code porting, and the setup of new workloads.

The test system initiates the warm-up phase for the next-generation LRZ supercomputer. LRZ's purchase explicitly includes the deployment of a test system and building out a training phase. The reason for this is the growing complexity of supercomputers, as well as the rapidly increasing costs. Initial investment and operation for Blue Lion will amount to €250 million, co-financed by the German Ministry of Research, Technology and Space (BMFTR) as well as the Bavarian Ministry for Science and the Arts (StMWK).

Since April 2025, LRZ's specialists have been meeting regularly with the HPE team: "Three topics are the focus of the preparations: code porting, energy efficiency, and workloads for AI," explained Gerald Mathias, head of LRZ's CXS (computational support) team. "With the help of the Blue Cubs, the planned system will be adapted for use and made fit for operations. "The test system will allow us to familiarize ourselves with how it works and set up administrative processes,", said Mathias.

During ISC25, NVIDIA had announced Blue Lion will feature Vera Rubin technology alongside HPE's next-gen supercomputing technology. These next-gen NVIDIA chips are specialized for HPC and artificial intelligence (AI) applications, but will not be available until 2026.

For this reason, the Blue Cubs are equipped with similar, already available Grace Hopper chips. The test system consists of eight Grace Hopper superchips in two nodes with four GPUs each. "The test system is the best equivalent to Blue Lion to-date and gives at least an idea of what the next supercomputer will bring," says Utz-Uwe Haus, head of HPE's Europe, Middle East, and Africa (EMEA) Research Labs, where systems for high-performance computing (HPC) in EMEA are planned. "It is intended to prepare administrators and support staff for what will be technically possible."

Focus on efficiency

While Intel and AMD's x86 architecture has dominated processors to-date, chips from ARM and NVIDIA offer more possibilities: GPUs are being better matched to CPUs, or even integrated into them. Blue Lion is designed to enable the combination of physical-mathematical and statistical models for simulations. Technology and usage are changing programming environments and workloads, and researchers must modify parts of their code so they are more specifically directed to the CPU, GPU, or storage. As a result, many algorithms and applications must be rewritten. To learn about implementation steps and functionalities or to clarify optimization requirements, the HPE and LRZ teams are experimenting with scientific programs on the test system. "It takes time for ported code to run

"Three topics are the focus of the preparations: code porting, energy efficiency, and workloads for AI."
Gerald Mathias, head of LRZ's CXS team

> **"The test system is the best equivalent to Blue Lion to-date and gives at least an idea of what the next supercomputer will bring."**
>
> Utz-Uwe Haus, Head of HPE's Europe, Middle East, and Africa (EMEA) Research Lab

efficiently on an HPC system, and intervention is often necessary," says Haus. "In addition, GPUs are energy-hungry, so if researchers don't use these processors efficiently, it substantially raises electricity costs."

Instead of runtime only, the focus in HPC is now on efficiency and, in turn, hardware control. System administrators use the Blue Cubs to determine how and when they will later reduce a processor's clock rate, for example, to reduce cooling requirements. The monitoring system used by the LRZ to control the work of computers and their environment provides guidance in this regard. "With this data, we can better adjust Blue Lion and observe which jobs require more cooling," said Haus. This experience will also help determine how many nodes later should be reserved for management. With the help of Blue Cubs, the teams are already developing procedures for system security, task allocation, and user identification: "We need to be able to rule out the possibility of Blue Lion being misused, for example for mining cryptocurrencies," said Haus.

The findings from the test phase will be incorporated into the future training program. "We can observe how much effort code porting requires and where support will be needed," explained Mathias. *Susanne Vieser*

Blue Lion is based on the Vera Rubin chip architecture.

Building a Virtual Training Academy for High-Performance Computing

The EuroHPC Virtual Training Academy (EVITA) is developing a standardised European training framework in HPC.

The EuroHPC Joint Undertaking (JU) has made major investments in supercomputing infrastructure, giving European science and industry powerful tools for remaining globally competitive. Maximizing the impact of these investments, however, will mean quickly growing a workforce with the skills and knowledge to use them effectively.

A new JU-funded project called EVITA (EuroHPC Virtual Training Academy) is one step toward this goal. Building on ongoing activities of the project HPC SPECTRA (see the spring 2024 issue of *InSiDE*), EVITA will establish a shared European framework for training in high-performance computing, including related fields such as artificial intelligence and quantum computing. In addition, the project will create an online repository of learning materials and make them available to training organisations and individual learners across Europe.

Coordinated by the Barcelona Supercomputing Center, EVITA brings together eight experienced centers for HPC training, including the High-Performance Computing Center Stuttgart (HLRS).

To establish this unified training framework, EVITA will develop a Competence and Qualification Framework (CQF), which defines guidelines on certification, learning pathways, professional profiles, and relationships between courses and models. Within the CQF, "skill tree" will structure learning pathways to address the needs of specific professional profiles, such as system administrator. Once the basic curriculum and learning paths have been defined, EVITA will begin developing training and learning materials suitable for HPC users across Europe, as well as a standardised template for developing new modules. Each module is a teaching/learning unit of 1-4 hours, including hands-on activities for learners to achieve learning objectives, with or without an instructor.

Using a cascade funding mechanism, EVITA will organise three open calls for proposals for module development over the course of the project. Creators of educational content will develop modules and courses based on pedagogical guidelines defined in the project. The resulting modules will be available on an open-source platform, enabling learners to gain valuable knowledge and skills. Moreover, EVITA will create a unified qualification and certification scheme in collaboration with the HPC Certification Forum. This scheme will both ensure consistency in course content, and provide learners with proof that they have developed specific competencies and skills.

EVITA is currently funded through March 2029, but ensuring the long-term sustainability of this shared training framework and its ability to evolve with new technological advances are also key objectives. The project will implement a governance structure and work with relevant stakeholders from across Europe to ensure that products developed in this Virtual Training Academy become integrated into the EU's digital education and training ecosystem, and continue to support the cultivation of new generations of skilled HPC professionals. *Christopher Williams*

PROJECT	EVITA
FUNDING AGENCY	Digital Europe Programme
FUNDING AMOUNT	Approximately €6 millon
START DATE	April 2025 – March 2029
PARTNERS	BSC, GWDG, TU Ostrava, Linköping University, Milano Polytechnic, TU Vienna, University of Luxembourg, HLRS

Scientists Release Record-Setting AI Earth Observation Model

Researchers at the Jülich Supercomputing Centre used the JUWELS Booster Supercomputer to develop TerraMind, an artificial intelligence foundation model that can be used for urban planning, improved agricultural outcomes, and other applications.

As the field of high-performance computing (HPC) enters an era defined by the rise of generative artificial intelligence (AI), scientists are using it to develop so-called foundation models that support research within a given scientific domain or research focus.

As their name implies, foundation models offer a basis for scientists to modify and build on for their own research goals. One example is the FAST-EO project: Funded by the European Space Agency (ESA) Φ-lab, FAST-EO has developed a foundation model for Earth observation tasks. The work could have far-reaching implications for understanding the impacts of climate change and for estimating its effects in Earth observation applications.

As a partner in FAST-EO, the Jülich Supercomputing Centre (JSC) at Forschungszentrum Jülich contributes both AI expertise and respective hardware with the center's JUWELS Booster supercomputer, to help develop the project's flagship model – TerraMind.

"With TerraMind, we hope to have an impact in agriculture, forestry, and urban planning, but also in things like civil protections regarding flooding or other natural disaster observation and response," said Dr. Rocco Sedona, researcher at JSC and collaborator on the project. "TerraMind helps us do Earth observation tasks much more rapidly, which helps inform people faster and with better information."

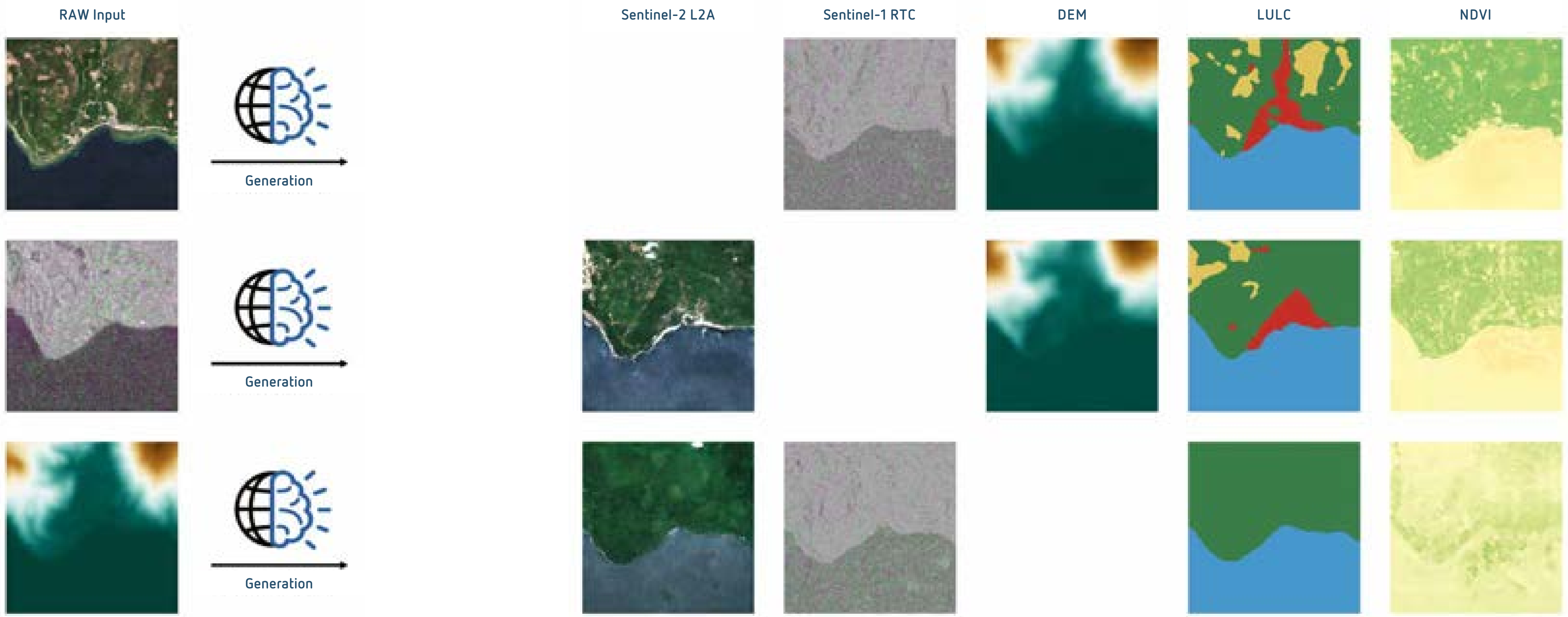
> "The GH200 superchips will be a big boost for working on this model."
> Dr. Rocco Sedona, JSC

HPC plays an essential role in foundation model training

Computer scientists primarily need two things to build a foundation model: a lot of data and a lot of computing power. This is because foundation models are essentially large-scale neural networks – a type of AI model that behaves similarly to how neurons exchange information in the human brain. They are trained with a large volume of data that starts off general and becomes increasingly specific. The resulting models can then be adapted easily for specific domains and tasks without the need for further large amounts of training data. Large language models such as ChatGPT were trained with hundreds of terabytes of written text, for instance. But for Earth observation, researchers use satellite imagery, elevation and other topographic data, and land-use maps. After compiling and organizing the relevant data, researchers give the model a task and provide feedback. This helps sharpen the model's ability to generate an image, text, or weather forecast.

Researchers currently train AI models using both pixel-level and token-level datasets. Just like with photography, pixels represent individual pieces of raw data; tokens, on the other hand, are larger chunks of abstract data that represent a part of an image, section of text, or another object of interest for a given AI model. "Tokens help us to say what is represented within a dataset," Sedona said. "In the case of Earth observation, that might be training the model to identify roads, airports, or even cornfields from satellite imagery, for instance."

TerraMind enables any-to-any generation across modalities, producing consistent outputs from diverse input combinations.



For TerraMind, the project collaborators used the JUWELS Booster supercomputer at JSC to train the model with over 500 billion tokens. Due to its GPU-heavy architecture, the JUWELS system is well-suited for such large-scale AI training. The team harnessed that potential to achieve top performance on the PANGAEA benchmark, a newly developed set of metrics for Earth observation models. PANGAEA evaluates a model's ability to correctly identify land cover, detect environmental changes, and perform other environmental monitoring tasks. TerraMind outperformed all previously evaluated models on PANGAEA.

"In order to train a model this large and complex, we need a resource like the JUWELS Booster," Sedona said. "The actual training uses a lot of computing resources, but we also have a significant amount of hyperparameter tuning. To generate 500 billion tokens, we gathered millions of raw images that needed to be pre-processed and post-processed, plus the time spent breaking them down into tokens. These tasks would be impossible without a machine like JUWELS."

Expanding the TerraMind foundation model supports larger research communities

While TerraMind has already outperformed other Earth observation models, the researchers have set their sights on expanding the model even further to improve performance and ultimately support even more research tasks.

JSC just inaugurated JUPITER, its newest flagship supercomputer (for more information on JUPITER's inauguration, visit page six). When fully operational, JUPITER will offer researchers roughly eight times as many GPUs as JUWELS. These GPUs are also tightly integrated with NVIDIA's Grace CPU as part of the company's GH200 Grace Hopper superchips.

"The GH200 superchips will be a big boost for working on this model," Sedona said. These superchips also have more memory, meaning that we will be able to use much more efficient algorithms than we currently have in place." While many AI researchers have significant experience using traditional large-scale HPC systems, others have limited experience using a system as powerful and complex as JUPITER. Sedona noted that while raw computing power is essential for the team to train a model like TerraMind, HPC centers like JSC also provide AI researchers with a variety of valuable services for training foundation models – JSC combines the expertise, raw computing power, and data infrastructure necessary to efficiently train a generative AI model. As part of JUPITER's installation, JSC has built a new Modular Data Centre (MDC), which houses dedicated storage systems and provides project collaborators with more efficient data management options.

Additionally, Sedona indicated that the project team wants to make TerraMind as open and easy-to-use as possible. "We believe in the principles of open science," he said. "This is publicly funded research, and we want to be open with what our model can achieve and make it available to as many people as possible." *Eric Gedenk*

The team has a pre-print publication available on arXiv.org: https://arxiv.org/pdf/2504.11171.

For more information about TerraMind, you can also visit the FAST-EO website, https://www.fast-eo.eu/.

PROJECT	FAST-EO
FUNDING AGENCY	European Space Agency (ESA)
FUNDING AMOUNT	€100,000
START DATE	January 2024 – September 2025
PARTNERS	European Space Agency (ESA), German Aerospace Center (DLR), Forschungszentrum Jülich, KP Labs, IBM Research

Five-Part Event Series Explores AI on HPC Systems for SMEs

Representatives from industry highlight successful applications of artificial intelligence for training large language models, managing energy usage, and improving manufacturing.

In collaboration with Baden-Württemberg networking and knowledge-transfer platform bwcon and artificial intelligence company Seedbox.ai, the High-Performance Computing Center Stuttgart (HLRS) launched a 5-part web series this summer focused on artificial intelligence on high-performance computing systems. With a focus on increasing understanding of the opportunities high-performance computing (HPC) and artificial intelligence (AI) offer for small and medium-sized enterprises (SMEs), the series brings together AI experts at HLRS and companies in Baden-Württemberg across multiple sectors. By highlighting successful applications, the ongoing series has been demonstrating how AI is changing engineering, manufacturing, and other industries.

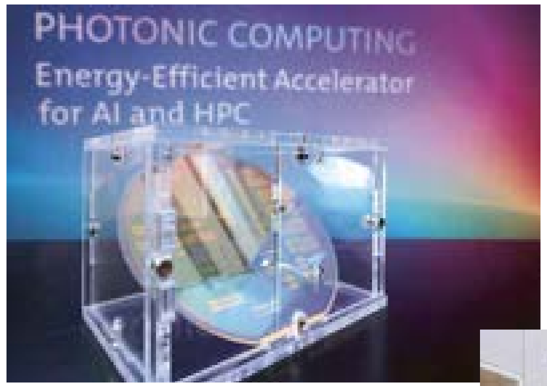
In the first episode Dennis Hoppe, head of HLRS's Department of Converged Computing, introduced HLRS and its capabilities for HPC and AI, including hybrid workflows that integrate the two. Dennis Dickmann, CEO of Seedbox.ai, followed by explaining why his company turned to HLRS and its Hunter supercomputer to train KafkaLM, a large language model incorporating 20 European languages. KafkaLM is not intended to replace more widely known LLMs like ChatGPT, Dickmann said, but because it runs at HLRS, it offers German companies a secure tool for the storage and analysis of proprietary data, ensuring compliance with data protection regulations such as GDPR. Dickmann highlighted KafkaLM's utility in data-intensive tasks in many business settings, such as managing knowledge across departments, evaluating and classifying documents, or generating content for reports, analyses, or internal communications.

The second episode looked at the challenges that industry can face in using modern energy systems, and how HPC and AI can make their operation more efficient. Matthias Heitmann, European Sales Director at Bosch Software and Digital Solutions, explained how AI models and sensor data can dynamically manage complex HVAC systems, reducing peak load and energy usage. His presentation considered the technical infrastructure Bosch has developed for this purpose, the savings potential of this approach, and how it can be reproduced at other locations.

Additional events in the series will focus on AI-driven manufacturing and real-time automation, and intelligent assistance for product development and services within the field of precision engineering. The final event will be held in conjunction with the inauguration of a new, AI-optimized, EuroHPC Joint Undertaking supercomputer that will be installed at HLRS in conjunction with the AI factory HammerHAI. The new machine is expected to arrive in 2026. *Christopher Williams*

Information about the series, including information about upcoming events and video of past events is available at: https://tinyurl.com/2w696kya

EVENTS



The wafer housing LRZ's new photonic server.

German Minister for Research, Technology and Space Dorthee Bär and Bavarian Minister for Science and the Arts Markus Blume tour LRZ to see the center's experimental photonic-processor-based system.

German Minister for Research, Technology and Space Dorothee Bär and Bavarian Minister of Science and the Arts Markus Blume tested the Q.ANT live image recognition demo.

A close-up of the Q.ANT chip.

The first installation of a photonic server in a real-live data center environment. The system consists of 3 servers and the technology will be evaluated by LRZ's Future Computing team.

Computing with Light

The Leibniz Supercomputing Centre just installed the world's first photonic AI computer at a data center.

A world premiere at the Leibniz Supercomputing Centre (LRZ): For the first time, a photonic computing system, the Native Processing Server (NPS) developed by Stuttgart-based deep-tech start-up Q.ANT, has been installed and put into operation in a data center. During an event in mid-July, Federal Research Minister Dorothee Bär and Bavarian Science Minister Markus Blume visited LRZ to welcome the system and gain insights into how this future technology works and the possibilities it offers researchers.

Less power, more performance – with light

Photonic chips compute directly with light and are well-suited for artificial intelligence (AI) methods, including image and pattern recognition. Q.ANT produces its processor based on lithium niobate, a new material, which enables highly efficient linear and nonlinear optical effects for computing. This allows the Q.ANT NPS to compute analog, faster, more precisely, and in parallel using various light wavelengths – and all at room temperature. Thanks to its high computing density and speed, the Q.ANT NPS is expected to deliver significantly more computing power for AI than traditional computers, work 50 times faster, and consume only one-thirtieth of the amount of electricity as it does not require active cooling.

The system currently operates with 16-bit precision and aims to deliver near 100 percent accuracy for all operations. The system is being seamlessly integrated into existing LRZ computing infrastructure and supports AI frameworks such as PyTorch, TensorFlow, and Keras. "For the first time, we are operating photonic processors in a high-performance computing center with real-world tasks," said Dr. Michael Förtsch, founder and CEO of Q.ANT. "This shows that light-based processors have made the leap from research to real-world application. This is a crucial step toward integrating photonic computing into mainstream computer architecture by 2030."

HPC architectures of the future

Together with LRZ, Q.ANT will optimize its system and explore whether it is suitable for hybrid digital-analog architectures for future supercomputers. The goal of the joint project is to develop benchmarks and practical use cases. "Photonic processors offer an innovative and promising way to accelerate AI and simulation workloads. This brings us closer to our goal of establishing ever more energy-efficient infrastructures for supercomputing and AI," explained Prof. Dr. Dieter Kranzlmüller, Head of the Board of Directors of LRZ. "The NPS from Q.ANT can be easily integrated into our existing infrastructure, allowing us to evaluate it immediately in practical scenarios."

For the evaluation, specially optimized Q.ANT NPS have been installed to test real applications, particularly in the areas of AI inference, computer vision, and physics simulations. Second and third-generation NPS systems will be deployed for more in-depth assessment over the next several months.

"The integration of Q.ANT's photonic processor into the Leibniz Supercomputing Centre is an impressive testament to German cutting-edge technology and a success story of German research funding," said Dorothee Bär, Federal Minister for Science, Technology and Space during the event. "We support groundbreaking innovations that strengthen our scientific leadership in the world and our technological sovereignty." The project is funded by the Federal Ministry for Research, Technology and Space.
Susanne Vieser

GCS Industry Day 2025: Supercomputing and AI Drive Private Sector Innovation

At GCS Industry Day 2025, JSC and GCS staff worked together with regional industry leaders to demonstrate how they can lower the barriers to entry for businesses to use high-performance computing and artificial intelligence to drive innovation and economic competitiveness.

High-performance computing (HPC) and artificial intelligence (AI) can accelerate innovation processes and strengthen the competitiveness of businesses. While some companies have been using these technologies for years, for many others, this is still uncharted territory.

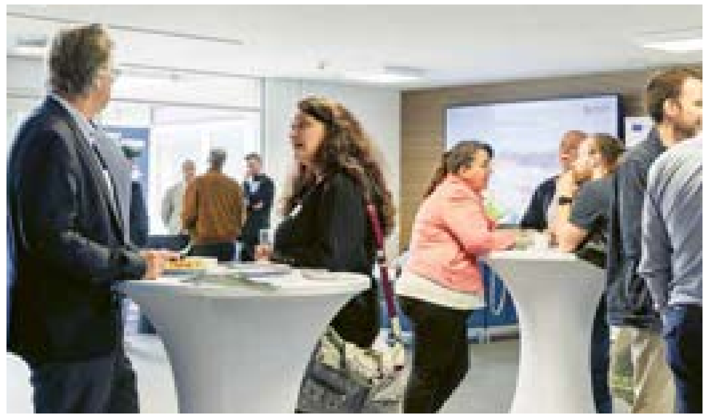
Bringing these digital future technologies closer to companies was the stated goal of the GCS Industry Day, in May 2025 at the Jülich Supercomputing Centre (JSC). The event showcased how companies can benefit from JSC's powerful infrastructure, its extensive expertise, and the latest funded projects in the fields of HPC and AI.

The event was organized by JSC in cooperation with the Gauss Centre for Supercomputing (GCS) and the Chambers of Commerce and Industry (CCI) of the Rhineland region: Aachen, Cologne and Mittlerer Niederrhein.

Real-world use cases and engaging discussions

At the heart of the event were presentations highlighting concrete business use cases. Representatives from Görtzen Stolbrink & Partner mbB, FEV Group GmbH, and IANUS Simulation GmbH shared first-hand experiences of using supercomputers and AI in their work. These included simulations of fire spread and evacuation movements used in building planning to ensure fire safety, as well as AI-assisted optimizations of flow simulations for product development and manufacturing.

The event offered the chance for a lively exchange between scientists and company representatives – during the moderated panel discussion, in discussion groups following the presentations, and in informal networking conversations.



GCS Industry Day 2025 participants enjoyed the various networking opportunities at the event.

EVENTS

These interactions fostered mutual understanding and laid the foundation for future collaborations.

To conclude the event, participants were invited to take guided tours of JSC, getting a closer look at and a greater understanding of Jülich's HPC infrastructure.

Lowering barriers to businesses to enter HPC

With its cutting-edge infrastructure and broad expertise, JSC offers optimal conditions for businesses to gain initial experience with HPC through low-threshold access. JSC's Industry Relations Team has been supporting businesses in getting started with supercomputing for over ten years.

Since 2022, the WestAI Service Center, in which the JSC is a key partner, has been offering extensive services for small and medium-sized enterprises, start-ups and academic researchers specifically focused on artificial intelligence. This includes expert consultations on AI concepts, the implementation of new AI use cases as prototypes, and access to computing resources, all provided free of charge.

By fall 2025, the JUPITER AI Factory in Jülich will become a central pillar of Europe's AI infrastructure. Start-ups, small and medium-sized enterprises, and large companies alike will gain access to Europe's first exascale-class supercomputer, JUPITER, for AI training purposes.

The GCS Industry Day 2025 highlighted the central role that supercomputing and AI play for national companies and demonstrated how to successfully build bridges between public research organizations and the private sector.

Previous collaborations between JSC and companies have primarily focused on industrial applications of HPC and AI. However, discussions at the event made clear that other sectors less familiar with HPC, such as logistics and the service industry, could also benefit from such partnerships in the future.

The panel's closing message was a clear call to action: HPC and AI offer tangible benefits for many businesses. Take the first step – be courageous!
Maximilian Tandi, Sohel Herff

Sohel Herff from the JSC Industry Relations Team speaks at GCS Industry Day 2025.

The event offered the chance for a lively exchange between scientists and company representatives during the panel discussion.

News Briefs

JSC

JSC Supports EduBrains Training Program on Future-Relevant E-Health Technologies for Brain Health Research

As part of the "EduBrains" training program, organized by EBRAINS Germany, around 40 students and doctoral candidates in STEM subjects will soon gain access to a specialized, interdisciplinary training and education program in the field of future-relevant e-health technologies for brain health research. An open call for applications was published in July 2025. The aim of "future e-health" is to promote highly qualified scientists at the interface of STEM and medicine, with an emphasis on information technology, data science, and artificial intelligence. As part of EduBrains, participants will get to know the European digital research infrastructure EBRAINS, learn new techniques and acquire new skills to address scientific and clinical questions in brain health research using state-of-the-art methods and FAIR (findable, accessible, interoperable, and replicable) data principles. The training program will familiarize students with the necessary theoretical, technical, and applied concepts required to create innovative research plans, develop new technologies and translate them into clinical solutions. In addition, the program includes training on ethics, dual use, data protection, and FAIR Open Science.

To read the full article, please visit: https://www.fz-juelich.de/en/jsc/news/news-items/news-flashes/2025/edubrains-starts

Helmholtz GPU Hackathon 2025 – Intensive Hacking Towards JUPITER

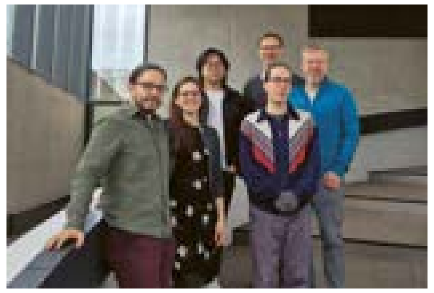
From April 1–11 2025, JSC hosted the 2025 Helmholtz GPU Hackathon. The event drew 52 researchers, developers, and data scientists from Helmholtz institutes and beyond. The event was organized by the Helmholtz Information and Data Science Academy (HIDA), NVIDIA, and the OpenACC Organization. The hackathon, now firmly established as a highlight of the Helmholtz calendar, focuses on optimizing complex high-performance computing (HPC) applications and increasingly integrates AI technologies. Seven teams from HPC and AI fields dedicated their time to optimizing application performance, with a particular focus on leveraging the NVIDIA GH200 superchip that powers JUPITER – Europe's first exascale system. While the construction site of the JUPITER system was just 50 meters from the hackathon venue, participants were able to access its initial hardware via the JEDI early access machine, which not only serves as a launch pad for JUPITER, but also remains the world's most energy-efficient supercomputer. The hackathon fostered both significant technical progress and valuable collaborations. As participants departed, they took with them both accomplishments and fresh ideas, setting the stage for continued development and innovation in HPC and AI across Helmholtz and its partners.

For more information, please visit: https://www.fz-juelich.de/en/jsc/news/news-items/news-flashes/2025/helmholtz-gpu-hackathon-2025-2013-intensive-hacking-towards-jupiter

JSC Involved in Three Clusters of Excellence

JSC is participating in three Clusters of Excellence under the Excellence Strategy of the German federal and state governments. The "Dynaverse" cluster addresses the central question of how the coupling of nonlinear physical processes influences the universe to the current day. Such processes control the dynamic evolution of the universe and operate on a wide range of timescales, from fractions of a second to billions of years. What is fundamentally new is that astrophysical research will be closely linked to the latest methods in data processing and artificial intelligence. The "Color meets Flavor" cluster is dedicated to the search for new phenomena in strong and weak interactions – a central challenge for the precision physics of the future. The focus is on the physics of quarks and the question of how they form complex binding states. At the same time, the properties of the Higgs boson will be investigated, and the search for the axion will be continued. The cluster, "ML4Q – Matter and Light for Quantum Computing," aims to create new computer and network architectures based on the principles of quantum mechanics. In the second funding phase of the Excellence Strategy, which begins in 2026, ML4Q will focus on the integration of quantum hardware and software. The Jülich UNified Infrastructure for Quantum computing (JUNIQ) at JSC will form one of the five pillars supporting ML4Q science initiatives. JSC's role in the Clusters of Excellence spans the development of scalable software frameworks, the integration of diverse simulation models, the optimisation of HPC resource usage, and addressing issues of long-term reproducibility as well as FAIR data practices (findable, accessible, interoperable, reusable).

For more information, please visit: https://www.fz-juelich.de/en/jsc/news/news-items/news-flashes/2025/jsc-involved-in-three-clusters-of-excellence



JSC Plays Key Role in Development of Open Multilingual LLMs for Germanic Languages

TrustLLM is an EU initiative focused on building open, trustworthy, and multilingual large language models (LLMs). At the second annual consortium meeting of TrustLLM, which ran June 11–13 2025, JSC joined partners across Europe at the University of Iceland. The event brought together leading voices in AI research, language technology, and policy to shape a European approach to inclusive artificial intelligence. Representing one of Europe's leading high-performance computing institutions, the JSC team contributed to discussions on the technical underpinnings of multilingual LLM development, data infrastructure, and evaluation frameworks that promote trust and transparency in language models. JSC is in the lead for work package six of the TrustLLM project, which aims to provide the technical foundation for large-scale training within the European HPC environment. JSC is pioneering advanced algorithms that significantly boost model capabilities while crucially reducing the computational footprint during both training and evaluation – laying the groundwork for scalable, sustainable AI in Europe.

For more information, please visit: https://www.fz-juelich.de/en/jsc/news/news-items/news-flashes/2025/trustllm-meeting-in-iceland-jsc-contributes-to-europe2019s-multilingual-ai-future

HLRS

Hunter in Top 3% for Energy Efficiency on Green500 List

In its inaugural appearance in the Green500 List, the Hunter supercomputer at the High-Performance Computing Center Stuttgart (HLRS) came in at position 12 worldwide. The Green500 List ranks all Top500 systems based on their HPL benchmark performance per watt of energy consumed, and Hunter achieved an energy performance of 64,653 gigaflops per watt. The system is based on the HPE Cray EX4000 platform from Hewlett Packard Enterprise. Its theoretical peak performance of 48.1 Petaflops is enabled by AMD Instinct MI300A accelerated processing units. At nearly double the speed of its predecessor, Hawk, Hunter uses 80% less energy at peak performance. This is supported by its use of 100% fanless direct liquid cooling system architecture developed at HPE, while a dynamic power capping method is being implemented to maximize the system's energy performance. Sustainability and energy efficiency have long been central components of HLRS's strategy. It remains the only HPC center of its size to be certified by the Eco-Management and

Audit Scheme, the world's most demanding standard for environmental performance. Planning for its next supercomputer, called Herder, and a new building called HLRS III are also being done with sustainability in mind. New infrastructure will capture waste heat generated by Herder and deliver it to buildings on the University of Stuttgart campus.

First Benchmark Study of the AMD MI300A APU for Training Large Language Models

While past benchmark studies have provided performance data of AMD's MI250X and MI300X processors, a paper presented at the May 2025 Cray User Group Technical Meeting offered the first rigorous benchmark study of the AMD Instinct MI300A accelerated processing unit (APU) for artificial intelligence (AI) applications. For high-performance computing (HPC) centers trying to decide whether to invest in a hybrid, APU-based system or in a separate, more specialized system optimized for machine learning and generative AI, the paper demonstrates that the MI300A is capable of achieving high performance in the training of large language models (LLMs). The publication also offers practical insights that system architects, AI practitioners, and HPC centers can use to leverage the full capabilities of the APU processor for deep learning and AI applications. The study resulted from a collaboration involving representatives of Seedbox.AI Lab, Hewlett Packard Enterprise (HPE), AMD, and HLRS. During this research, a data compression approach being developed by Seedbox.AI demonstrated the viability of sparse training models as an alternative to large models. The team completed its experiments on HLRS's Hunter supercomputer.

HLRS Conference Explores Uncertainty in Simulation

Uncertainty is an inherent part of simulation, whether in predicting weather, identifying new drugs to treat disease, or optimizing the design of a new engine. Data availability, modeling choices, and randomness, for example, are all factors that affect simulation results, while machine learning and artificial intelligence – so-called "black box" technologies – give no clues about how an algorithm arrives at its result. Uncertainty is not just a concept in high-performance computing, of course. Other disciplines including economics, politics, philosophy, and science all have long traditions of reckoning with it in different ways. A recent conference held at the the High-Performance Computing Center Stuttgart (HLRS) sought to draw on such multidisciplinary expertise, exploring how dialogue between the humanities, social sciences, and computational sciences could illuminate the concept of uncertainty in new ways. Talks considered topics such as the causes of uncertainty in simulation, ways to improve uncertainty quantification, considerations for understanding and explaining uncertainties in policy-making contexts, and productive uses of uncertainty. The three-day international conference was the latest in the series "The Science and Art of Simulation" (SAS25), an annual meeting at HLRS organized by the Department of Philosophy of Computational Sciences that in the past has explored topics related to interpretability, trustworthiness, and uses of the simulation sciences in policy making. A forthcoming volume published by Springer Verlag will collect the conference proceedings.

HLRS Joins German AI Association

The High-Performance Computing Center Stuttgart (HLRS) has become a member of the German AI Association, the country's largest network for companies working in the rapidly evolving field of artificial intelligence. The German AI Association offers a platform for information sharing, network building, and raising the visibility of German innovation using artificial intelligence. It focuses on helping AI-focused businesses to address key challenges they face, and serves as a voice for the AI community on the national political level. "We are delighted to join the German AI Association," said Dr. Bastian Koller, Managing Director of HLRS. "This is a perfect forum for us to contribute our many years of experience in providing computing resources and support services for industry (including research and manufacturing) and science. At the same time, we look forward to interactions within the German AI Association that will enable us to expand our knowledge and our portfolio continually, both as HLRS and as coordinator of the AI Factory, HammerHAI."

Extending sys-sage to Integrate Quantum Hardware

A collaboration between the Technical University of Munich (TUM) and LRZ has led to the enhancement of the sys-sage library, a tool traditionally used in high-performance computing (HPC). This extension enables sys-sage to analyze and optimize quantum computing systems, facilitating their integration into supercomputing environments. The team's work was recognized with the Hans Meuer Award at ISC25 in Hamburg. Sys-sage processes data from supercomputers such as processor performance and data transfer metrics, helping manage and improve system operations. With the rise of heterogeneous computing systems, including quantum computers, sys-sage has been adapted to handle the unique characteristics of quantum hardware. This includes assessing qubit quality and filtering out noisy connections to ensure accurate computations. The extended sys-sage is a component of the Munich Quantum Software Stack (MQSS), developed under the Q-DESSI project within the Munich Quantum Valley initiative.

More information: https://tiny.badw.de/zdAtr8

LRZ Develops 10B Language Model

Do-you want to translate "Hock di hera, samma mera"* into German or English? Now there's Llama-GENBA-10B. The trilingual language model is based on Meta's Large Language Model (LLM) Llama, version 3.1-8B, and was trained by researchers at LRZ and Cerebras Systems with 10 billion parameters using a dataset of 164 billion tokens. Llama-GENBA-10B is an inclusive and resource-efficient base model that not only translates but also generates texts in English, German, and Bavarian. "Our model demonstrates efficient multilingual training on the Cerebras CS-2 system," explains Michael Hoffmann of the LRZ Big Data & Artificial Intelligence (BDAI) team. "To train Llama GENBA 10B, the CS2 system consumed around 35 megawatt hours of energy in 66 days." The group has just published a paper (preprint) on the method and challenges of training Llama-GENBA-10B. It compares the model's performance with other language models, such as Apertus-8B, gemma-2-9b and EuroLLM-9B. "In addition to performance, it was important to us to work with non-English data and, above all, with a dialect," says Jophin John from the LRZ BDAI team. Since most LLMs focus on English, Llama-GENBA-10B strengthens the preservation of less common languages and regional dialects, ultimately providing a blueprint for similar models that even small research teams can implement.

* *"Sit here, then there will be more of us."*

More information: https://www.lrz.de/en/news/detail/lrz-develops-10b-language-model

A Substation for LRZ

To continue providing science and research with powerful computing resources and supercomputers in the future, LRZ needs more energy. In the future, this will also come from the substation that the network operator Bayernwerk, a subsidiary of the utility company EON, is now building in the north of Garching, specifically designed to meet LRZ's requirements. The symbolic groundbreaking ceremony took place on October 6, and the substation is scheduled to go into operation in 2028.

Two transformers will each provide a capacity of 50 megavolt amperes (MVA) to ensure a redundant power supply for the LRZ. In mathematical terms, this corresponds to approximately 10,000 average photovoltaic systems on single-family homes, each with an output of ten kilowatts.

This gives the LRZ greater certainty when planning its next high-performance IT infrastructures. "Simulations of natural phenomena, modeling of climate data, artificial intelligence methods – science relies on increasingly powerful supercomputers and AI machines. Our electricity requirements will increase significantly in the coming years – not even the extremely energy-efficient operation of our data center can prevent that," said Prof. Dr. Helmut Reiser, deputy director of the LRZ. "The new substation creates the conditions for us to operate our IT infrastructures and thus makes our data center fit for the future."

More information can be found in the press release from Bayernwerk (German only): https://www.bayernwerk.de/de/ueber-uns/newsroom.html#/pressreleases/neues-umspannwerk-in-garching-energie-fuer-die-spitzenforschung-3408845

Replacing AI Training with Math Methods

Prof. Felix Dietrich from the Technical University of Munich (TUM) is exploring a novel approach to artificial intelligence (AI) by moving away from traditional training methods. Instead of relying on large datasets and extensive training processes, he and his team are developing AI models that utilize mathematical methods to directly interpret raw data. This shift aims to enhance efficiency and reduce the computational resources required for AI development. By integrating mathematical techniques, these models can potentially offer more accurate and faster analyses, opening new avenues for AI applications across various fields.

More information: https://tiny.badw.de/gO6jQ0

In his role at LRZ, Ajay Navilarekal Rajgopal is able to fuse his interests in mathematics, engineering, and artificial intelligence by supporting a diverse user base.

Staff Spotlight: From Auto Research to AI

Ajay Navilarekal Rajgopal, LRZ

Growing up, Ajay Navilarekal Rajgopal wanted to fuse his passion for mathematics with the excitement he felt watching vehicles at work – be they bicycles, trains, cars, or planes. He decided to combine these interests in a bachelor's degree in mechanical engineering at India's National Institute of Engineering. During his studies, he was introduced to courses on computational fluid mechanics that awakened a third passion to include in his career goals: he developed a burgeoning interest in computational science and the use of high-performance computing (HPC) for solving the world's most challenging scientific and engineering problems. "When I was introduced to more advanced coding and realized how much math was involved, I saw computational engineering as a way to fuse all of my interests together," he said.

Ajay started working for Mercedes Benz Research and Development India, which not only gave him first-hand experience practicing his craft, but also exposed him to German work culture and offered him the chance to start learning the German language. He parlayed that experience into a graduate program in Germany, earning his master's degree in computational sciences and engineering at the Technische Universität Braunschweig (TU Braunschweig).

As he got deeper into the field, Ajay realized that the skill set he had been developing was broadly applicable within the greater HPC ecosystem. "In a way, I realized that the core principles of R&D are what I would call domain agnostic," he said. "To me, that means that once you start to think about what you are actually doing with a code base, you realize that many research domains are all trying to do the same thing: you want to improve the efficiency of codes so they will go faster or use less energy without sacrificing the quality of your simulations," he said.

As Ajay grew his HPC knowledge, the field was also in the process of major change. Specifically, artificial intelligence (AI) – a technology that had long been developed and incubated at public HPC centers – was coming into its own, and the rise of generative AI promised to radically alter how researchers could use large-scale computing resources to accelerate their research.

> "When I was introduced to more advanced coding and realized how much math was involved, I saw computational engineering as a way to fuse all of my interests together."
> Ajay Navilarekal Rajgopal, LRZ

"This was my second major eureka moment, and I realized that if you program neural networks well, they can make accurate predictions based on what you teach it."

Ajay Navilarekal Rajgopal, LRZ

Ajay was introduced to neural networks during his graduate studies and saw the power in using these methods to further accelerate computational research. Neural networks operate similarly to how the human nervous system processes information and form the foundation for large-scale AI model development. "This was my second major eureka moment, and I realized that if you program neural networks well, they can make accurate predictions based on what you teach it," he said. "We were free to choose our own research subjects, and I decided that I wanted to focus more on machine learning as the core of my graduate studies."

Working as a user support specialist at LRZ, Ajay's work focuses on the intersection of his passions, but he also gets to support research teams at an institution solely focused on scientific outcomes. "One of the best parts of working at LRZ is that we are funded by the state," he said. "As a result, we are not interested in researchers' money, or trying to monetize researchers' data – we are only providing a service that is otherwise difficult for many research institutions to build or maintain on their own."

In his role, Ajay sees supporting users in the AI era as a two-pronged issue: LRZ staff must help researchers with less experience using HPC or other large-scale computing resources learn to use these systems efficiently while also supporting long-time HPC users in adapting their workflows to take advantage of AI. He pointed out that many engineering research projects rely on legacy codes that have been collaboratively developed over years. Accordingly, he sees a core part of his role as helping users develop AI models that can be integrated into these long-running research tools.

Since LRZ is one of the three computing centers that comprise the Gauss Centre for Supercomputing (GCS), Ajay also benefits from a larger network of computational science experts for knowledge sharing, collective problem solving, and sharing best practices. Together with other user support specialists at the High-Performance Computing Center Stuttgart and Jülich Supercomputing Centre, he feels motivated and well-positioned to deliver on the core GCS mission – supporting researchers at all levels in solving the pressing scientific and engineering challenges of today and tomorrow.

"One of my key responsibilities is providing access to these valuable resources," he said. "Even if a person is running a relatively simple algorithm but has a bigger idea they know cannot run on a laptop or smaller compute cluster, we want to be able to help them get started. I want to foster a close loop between researchers and our team, so that we're not just waiting for the user to come to us with a problem, but we are investing the time in helping develop solutions that can proactively benefit teams and, in some cases, the larger research community."
Eric Gedenk

Ajay considers working at a public science organization a benefit. "One of the best parts of working at LRZ is that we are funded by the state. As a result, we are not interested in researchers' money or trying to monetize researchers' data — we are only providing a service that is otherwise difficult for many research organizations to build or maintain on their own," he said.

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, and coordinates the EuroHPC JU AI Factory HammerHAI. 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.

HLRS
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 JUPITER 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.

CONTACT

Jülich Supercomputing Centre (JSC)
Forschungszentrum Jülich

Prof. Dr. Dr. Thomas Lippert
Phone +49 - 24 61 - 61 - 64 02 / th.lippert@fz-juelich.de

Prof. Dr. Kristel Michielsen
Phone +49 - 24 61 - 61 - 25 24 / k.michielsen@fz-juelich.de

Wilhelm-Johnen-Straße, 52425 Jülich, Germany
www.fz-juelich.de/jsc

Compute servers currently operated by JSC

System	Size	Peak Performance (Tflop/s)	Purpose	User Community
JUPITER BOOSTER	125 racks, 5,884 nodes Eviden BullSequana XH3000 23,536 NVIDIA GH200 Superchips 1,694,592 CPU cores NVIDIA Grace 4,964 TB memory (LPDDR5+HBM3)	1,258,140	Capability Computing and AI	European and German Universities, Research Institutes, and Industry
JSC CLOUD	72 nodes, 11,760 cores AMD EPYC & Intel NVIDIA V100/A100/RTX A6000/L40s/H100, AMD MI210 36 TByte memory 145 TByte NVM	760	Cloud Computing (IaaS and PaaS)	European and German Research Institutes
MODULAR SUPERCOMPUTER "JUWELS"	Cluster (Eviden): 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 (Eviden): 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 (Eviden): 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
	JURECA-HWAI: 32 nodes, 2,048 cores Intel Sapphire Rapids 8462Y 128 NVIDIA H100 GPUs 16 TByte memory	8,756	AI	German Research Institutes and Industry HAICORE3/WestAI
HPC/ CLOUD 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
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 receive InSiDE magazine twice a year. You can also request to receive the magazine as a PDF via email.

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, Vincent Miechielsen; www.groothuis.de

Print:
Buch- und Offsetdruckerei H. Heenemann GmbH & Co. KG

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

RECYCLED Paper made from recycled material FSC® C020935

Druckprodukt mit finanziellem Klimabeitrag ClimatePartner.com/53170-2510-1009

Contributing authors: Eric Gedenk, GCS, eric.gedenk@lrz.de; Christopher Williams, HLRS, williams@hlrs.de; Susanne Vieser, LRZ, susanne.vieser@lrz.de; Lisa Maiburg, JSC, l.maiburg@fz-juelich.de; Maximillian Tandi, JSC, m.tandi@fz-juelich.de; Sohel Herff, JSC, s.herff@fz-juelich.de; Stefan Krieg, JSC, s.krieg@fz-juelich.de; Kristel Michielsen, JSC, k.michielsen@fz-juelich.de; Susanne Pfalzner, JSC, s.pfalzner@fz-juelich.de

Photo credits:
Cover page 1: Philipp Deus/unsplash.com; pg. 2: Ethan Grey/unsplash.com; pg. 3: Alessandro Podo, Sascha Kreklau, HLRS; pg. 6: Forschungszentrum Jülich/Kurt Steinhausen; pg. 9: LRZ; pg. 10: Naumann for LRZ; pg. 13: HLRS; pg. 14–17: Institute of Aerodynamics and Gas Dynamics, University of Stuttgart. (Or IAG, University of Stuttgart); pg. 19: Marcel Aach; pg. 21: Salvatore Cielo; pg. 23: Salvatore Cielo and James Beattie; pg. 24: NVIDIA, Codioful/unsplasch.com; pg. 26: LRZ; pg. 28/29: Forschungszentrum Jülich, and the ESA Φ-Lab; pg. 32: LRZ/Veronika Hohenegger; pg. 34/35: Jenö Gellinek; pg. 39/37: JSC, Jülich Supercomputing Centre; pg. 38: ISC High Performance | Philipp Loeper; pg. 40–43: Veronika Hohenegger for LRZ; pg. 44: HLRS/Julian Holzwarth; 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