Deutsches Forschungszentrum für Künstliche Intelligenz
German Research Center for Artificial Intelligence

dfki ai next

How AI orients to the real world

From perception to action

AI-generated

AI

Human-Centric AI

Intelligent solutions for the knowledge society

DFKI has more than 35 years of experience in human-centric AI, conducting research in key areas of basic research and future-oriented applications, while always maintaining a focus on social relevance and scientific excellence in the field of artificial intelligence.

The German Research Center for Artificial Intelligence (DFKI) was founded in 1988 as a non-profit public-private partnership. It operates facilities in Kaiserslautern, Saarbrücken, Bremen, Lower Saxony (Osnabrück and Oldenburg), Darmstadt, with laboratories in Berlin and Lübeck, and a branch office in Trier.

It supports 29 research areas, thirteen competence centers, and eight living labs. Based on application-oriented research, it develops product functions, prototypes, and patentable solutions in the information and communication technology (ICT) sector. Funding is provided by public funding sources and industrial development contracts.

Project results and milestones are periodically reviewed institutionally and by a Scientific Advisory Board, an international panel of experts. In addition to the German Federal Ministry of Research, Technology and Space (BMFTR) and the federal states of Rhineland-Palatinate, Saarland, Bremen, Lower Saxony and Hesse, several leading German and international high-tech companies from a wide range of industries also have seats on the board.

www.dfki.de

Think about AI in Federal and European terms

International cooperation, scientific visibility, and strategic sovereignty

Advances in artificial intelligence are not made in isolation. The work of CERTAIN and other collaborations with strong European and international partners such as DFKI and INRIA show that scientific excellence is achieved when research is brought together across institute and national boundaries.

The network of German AI Competence Centers serves to strengthen the international reputation of German AI research while helping to establish AI made in Europe as a hallmark of excellence and trustworthy AI. Associations like the Robotics Institute of Germany (RIG) pool expertise, foster connectivity between disciplines, and demonstrate that AI research is often at its best where perception, orientation, and action are considered holistically, not separately.

The research now being discussed on major international stages – from CVPR, ICML, and other IEEE conferences to IJCAI – is shaping the next major advance in AI.

Germany's High-Tech Agenda establishes the political framework. The goal is clear: strengthen technological sovereignty, improve the accessibility of AI resources, and ensure that Germany remains a global leader for the next generation of AI.

This issue treats that as its starting point. It demonstrate how excellent research, international cooperation, and domain-specific networking can give rise to an outstanding AI landscape.

How AI orients to the real world

From perception to action

Artificial intelligence has reached a point where performance improvements alone are no longer enough. The crucial question is not how systems can become even bigger, faster, or more ubiquitous. The crucial question is what kind of AI Europe wants to develop to ensure future scientific autonomy, societal reliability, and technological agency.

This is the true challenge researchers face today: not merely to make AI more powerful, but to enable it to navigate the complexities of an open and human-centric world.

Germany and Europe are well-positioned for this challenge: scientific excellence, growing political awareness, and a public ever more alert to the fact that AI is not some distant future technology, but is already a critical structural issue for economic, scientific, and societal success. Consequently, now is the time to translate these insights into concrete goals. Where AI is viewed merely as a global race for scale, it will ultimately be subject to the development logic of others. However, those who take AI seriously as a research subject will be able to set their own standards – regarding transparency, energy efficiency, controls, and real-world integration. Technological sovereignty does not equate with market success, but with the ability to determine the foundation and direction of future systems.

Europe must invest precisely in these areas when developing the next generation of AI if it wishes to reduce dependence on closed, non-European systems. This includes new computer architectures, energy-efficient processing paradigms, and infrastructures – not designed for maximum scale, but for resilience, security, efficiency, and controllability (scientific and industrial).

Even more fundamental is the question of which model of AI we actually want to continue developing. The mounting success of large generative systems is undeniable. Equally undeniable, however, are their limitations: enormous data requirements, opaque decision-making processes, high energy consumption, and significant levels of unreliability in safety-critical or highly specialized systems. Many applications in industry, medicine, mobility, or the energy sector demand more than a system that merely appears powerful. It must be transparent, verifiable, and capable of explaining its decisions in robust contexts. This is precisely why one field of research is gaining traction, because it represents much more for Europe than just another technical alternative: hybrid, neuro-explicit AI. It combines learning methods with symbolic, knowledge-based, and analytical components.

This aligns nicely with the shift away from the simplistic “bigger is better” mindset. Small, specialized language models and frugal AI architectures demonstrate that efficiency itself can be a step forward. Such models require less energy, less special hardware, and can be

tailored more precisely to specific domains – from industrial processes to scientific data spaces or sensitive knowledge environments. Especially for a research and economic landscape that relies heavily on highly specialized, regulated, and SME-dominated systems, this is not a minor consideration, but a strategic advantage. Reliability under real-world constraints is more important than maximum generality at any price.

Add to this another ongoing shift: AI no longer generates only textual and visual content; it can now plan, coordinate, and manage expanded workloads. The focus is shifting once again with the transition to agents. As systems enter contexts of related actions, transparency, intervention options, and human oversight become key.

The next step in AI research is not to unleash the next level of autonomy, but rather to ensure the societal and technical integration of systems that are controllable and transparent. Only then does raw computing power transform into reliable agency.

These challenges for technology and science converge in this issue of ai next. The real question for the future is not only what AI can recognize but also how it orients itself to real-world, dynamic, and ambiguous environments. How does perception transform to reliable knowledge? How is knowledge used to create a robust framework of classification? And how to take action that is not merely efficient, but also justifiable? This move from perception to orientation to action presents not only a plausible narrative theme for this publication, but also a precise research horizon for the coming years.

Europe has no need to copy foreign models in pursuit of this horizon. It can rely on its own strength in developing AI research centered on reliability, clarity of orientation, and real-world integration – open to international collaboration yet steadfast in its own scientific standards. The next step will be defined not by the loudest system, but by the one that combines transparency, efficiency, and accountability with performance. That is precisely where AI begins to operate in the real world, not within isolated data spaces.

« The promise of hybrid, neuro-explicit AI lies in scaling up existing models, rather than in merely increasing their size, as it involves a synergistic concept of machine intelligence that combines the strengths of probabilistic and deductive AI approaches. Systems of this kind do more than recognize patterns: they explicitly represent knowledge, structure their decisions, and allow auditing under defined conditions. This is a scientific view that does not pit performance against trustworthiness but systematically brings these two together. »

Prof. Dr. Antonio Krüger, CEO DFKI

Papers and authors:

From raw geometry to meaningful scenes

Three approaches to 3D scene understanding, attention, and reconstruction

Understanding 3D scenes begins where geometry, relationships, and perspective meet. Researchers at DFKI recognized this dynamic and made it a priority of their work, presenting their findings at the June Conference on Computer Vision and Pattern Recognition (CVPR) 2026: Spatial reconstruction is no longer viewed as the end goal, but rather as a starting point for semantic interpretation, human-centric orientation, and targeted intervention within complex 360-degree environments.

ReLaGS represents a fundamental advance in 3D scene understanding. The method combines Gaussian splatting with hierarchical scene representation and relational logic in a way that enables machines to capture and model the spatial and semantic relationships among objects. The process turns a raw reconstruction into a meaningful scene.

Spatial constructs, part-whole relationships, and language descriptions of relationships are explicitly represented by geometric points and surfaces in a structured

environment. ReLaGS achieves a scientific innovation by joining hierarchical semantics and relational 3D representation in a unified, scene-agnostic approach. This approach creates an open 3D Scene Graph without the need for scene-specific training, while at the same time improving spatial and semantic stability through pruning and robust attribute aggregation. In a world where structured 3D models of environments are gaining international importance, experts are teaching machines not only what to recognize in a scene but also the relationships among objects. DriverGaze360 explores this issue in another, equally important direction: How does a person look at a complex scene?

DFKI researchers have compiled a dataset of approximately one million gaze-annotated frames from a realistic 360-degree driving simulator to demonstrate how driver attention can be captured under controlled yet realistic conditions. Such measurements are relevant not only for analyzing human perception, but also for assistance systems that need to understand what drivers are actually focusing on at any given moment. This is lacking in previous research where traditional datasets primarily capture frontal views and underestimate lateral or rearward attention dynamics, for example, when changing lanes, merging onto a highway, or checking behind the vehicle.

DriverGaze360 is more than just a topic related to traffic. The work demonstrates that scene understanding is not merely a matter of machine segmentation and geometric relationships, but also involves situated, human orientation under realistic conditions. Tracking additional objects is key to the system's ability to learn not only gaze distribution, but also which object is the focus of attention at any critical moment. As a result, attention becomes semantically connectable and describable as a structured relationship between the person and their environment.

The third approach, as introduced by Inpaint360GS, is active scene editing. The core challenge of any realistic 3D representation is that scenes are often incomplete, obscured, or have gaps. Our method addresses this by filling in missing scene and spatial elements in 360-degree environments in an object-aware manner. What matters most here is how something is reconstructed in the scene. Inpaint360GS fills in missing areas in a way that preserves object logic, spatial structure, and semantic coherence.

The overarching scientific link to the ReLaGS and DriverGaze360 projects is clear. While ReLaGS demonstrates how scenes become interpretable and DriverGaze360 reveals how attention is distributed within them, Inpaint360GS shows how, on this basis, it is possible to intervene in scenes without destroying their internal structure. Perception does not automatically lead to action, but it does provide the basis for editable nodes that integrate geometric accuracy, semantic understanding, and contextual awareness.

The three studies together demonstrate more than a methodological improvement. They mark an advance toward AI systems that are able not only to reconstruct spatial environments, but also to understand and process their structure, relationships, and gaps. By integrating scenes, attention, and reconstruction, the foundation is set for reliable robots, more robust assistance systems, more precise digital twins, and visual AI that does not settle for what is simply visible but also understands context.

CVPR 2026: A Stage for Visual AI

The IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) is one of the most important international conferences in computer vision research. This year‘s conference took place in Denver from June 3-7, 2026. For DFKI, it was a key event to discuss current work on 3D scene understanding, relational logic, attention modeling, and object-aware reconstruction. This issue features three contributions from DFKI‘s Augmented Reality Research Department: ReLaGS, DriverGaze360, and Inpaint360GS. Together, they show how geometry, attention, and scene editing can be viewed as part of a related research movement.

Papers and authors:

The person depicted is fictional and AI-generated.
Microscopy: Louisa Howard, Public Domain.

The bottleneck in microscopy

How AI foundation models are reshaping image analysis in the life sciences

Modern microscopes provide image data at resolutions and volumes that are revolutionizing biological research. Consequently, the bottlenecks are encountered less and less at image acquisition and more often with the analysis: cells, nuclei, and organelles must be precisely delineated and consistently annotated across large datasets. Segmentation is what turns microscopic images into reliable research data.

A recent study published in Nature Methods, “Segment Anything for Microscopy,” investigates whether a foundation model developed for everyday images can be adapted to microscopic data without losing its flexibility. Researchers succeeded in demonstrating that the „Segment Anything Model“ (µSAM) can be adapted for light and electron microscopes to more reliably capture biological structures while simultaneously supporting both interactive and automated segmentation tasks.

The scientific significance of this work is found not only in the improved results. A crucial aspect is adapting a generated AI image model to a domain where the visual logic differs fundamentally from that of the everyday visual world. Object boundaries, contrasts, shapes, and spatial relationships follow different rules in microscopy; it is particularly important to determine the true extent of modern AI’s generalization capabilities.

« The µSAM project demonstrates the productive use of foundation models in highly specialized scientific imaging domains – provided they are methodically and consistently adapted to the specific domain. Based on this experience, we are preparing a spin-off to perform cell culture analysis that should effectively translate the potential of this research into a concrete application. »

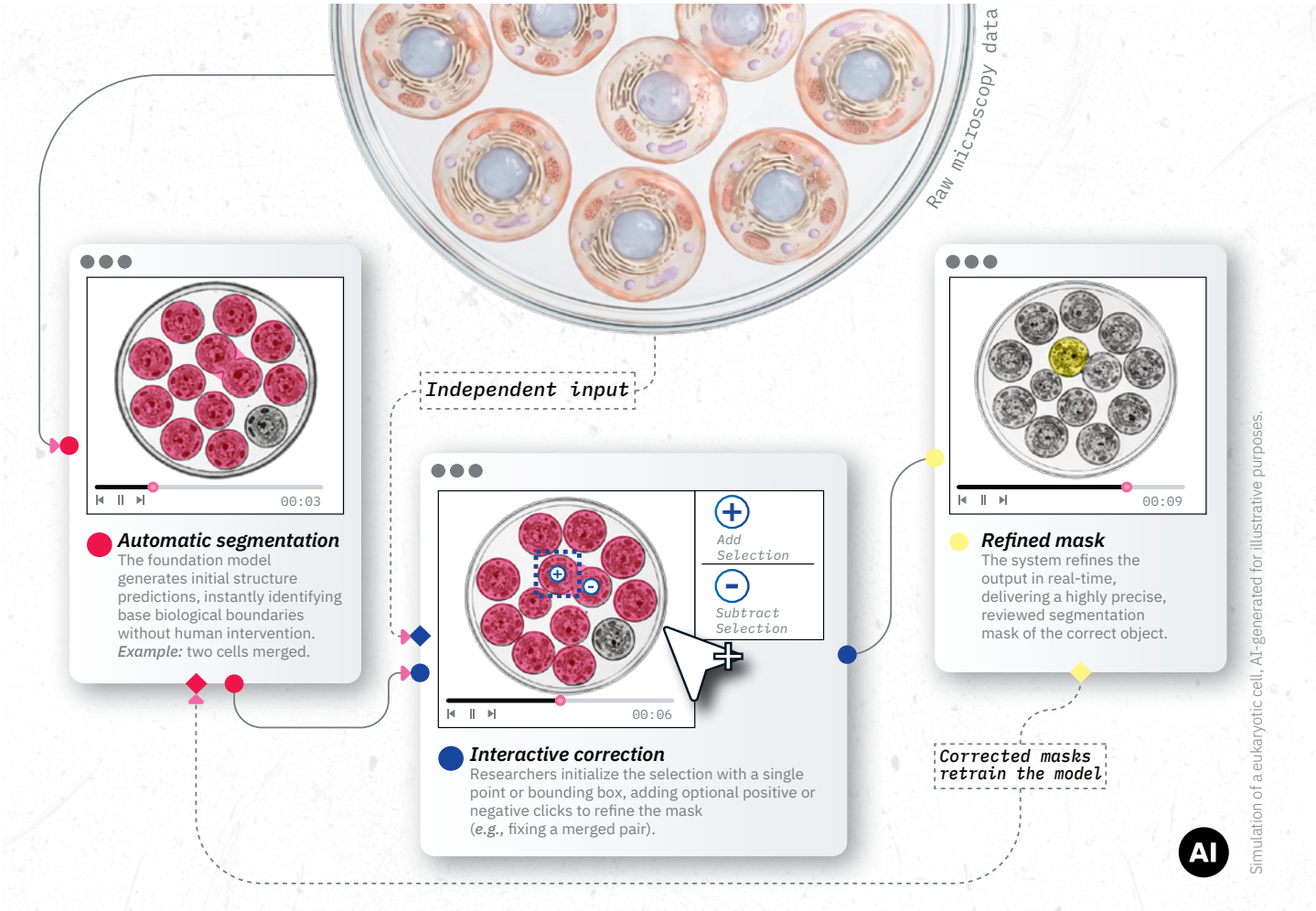
Prof. Dr. Andreas Dengel,
Head of Smart Data & Knowledge Services Research Department, Executive Director, DFKI Kaiserslautern

This approach is especially relevant where human expertise and model performance intersect. µSAM is designed to be an interactive working environment: researchers provide a few prompts; the system fills in the contours and incorporates corrections into the ongoing workflow.

µSAM transforms a slow, routine task into a collaborative workflow that accelerates annotation without sacrificing scientific oversight.

For the life sciences, this is more than just a technological innovation. Fields such as cell biology, cancer research, and developmental biology generate vast amounts of data that can only be analyzed manually at considerable cost in terms of time and effort. The research results demonstrate how this bottleneck can be overcome, not by replacing scientific work, but through a more precise integration of the models, interfaces, and domain knowledge.

This study also contributes to the international debate surrounding foundation models. It states that the case for such systems lies not in abstract, universal applicability but in their precise translation into actual scientific applications. That is where it is determined whether it is just an exciting new model or whether it delivers a genuine impact on scientific practice.



Papers and authors:

Trained only in English, AI will remain exclusive

Synthetic training data helps to create models for low-resource languages

Large language models (LLMs) perform extremely well in those languages for which a wealth of digital data is available. Vast quantities of text are available for high-resource languages like English and German. However, many other languages are barely represented in the training of modern AI systems. Currently, there are no high-performance AI applications available for many of the minority language communities. This disparity leaves the potential economic and social benefits of modern language technologies largely untapped.

Can synthetically generated training data bridge the gap? A research team from DFKI, the Brno University of Technology, and the Kempelen Institute of Intelligent Technologies in Bratislava thinks they have the answer. Their work explored the central issue: under what conditions can large language models (LLMs) generate sufficient high-quality training data to be suitable for training additional AI models?

The researchers used four open-source LLMs of varying sizes, including models from the Llama and Gemma families. These models generated training data for eleven typologically diverse languages, ranging from high-resource languages such as English and German to languages with a significantly smaller digital presence, such as Welsh, Romanian, Azerbaijani, Slovenian, and Telugu. The team then used the generated data in their investigation of three key language processing tasks: user intent recognition, topic classification, and sentiment analysis (positive or negative opinions).

Researchers trained new AI models to verify the quality of the synthetically generated data. The results were compared to results from models trained on real, human-annotated “gold standard” data. In this way, a precise determination was made of how reliable the synthetic datasets actually are.

The team found that the performance gap to the models based on real training data had significantly narrowed. The new models nearly matched the performance levels achieved with the gold standard data, particularly in the task of user intent recognition.

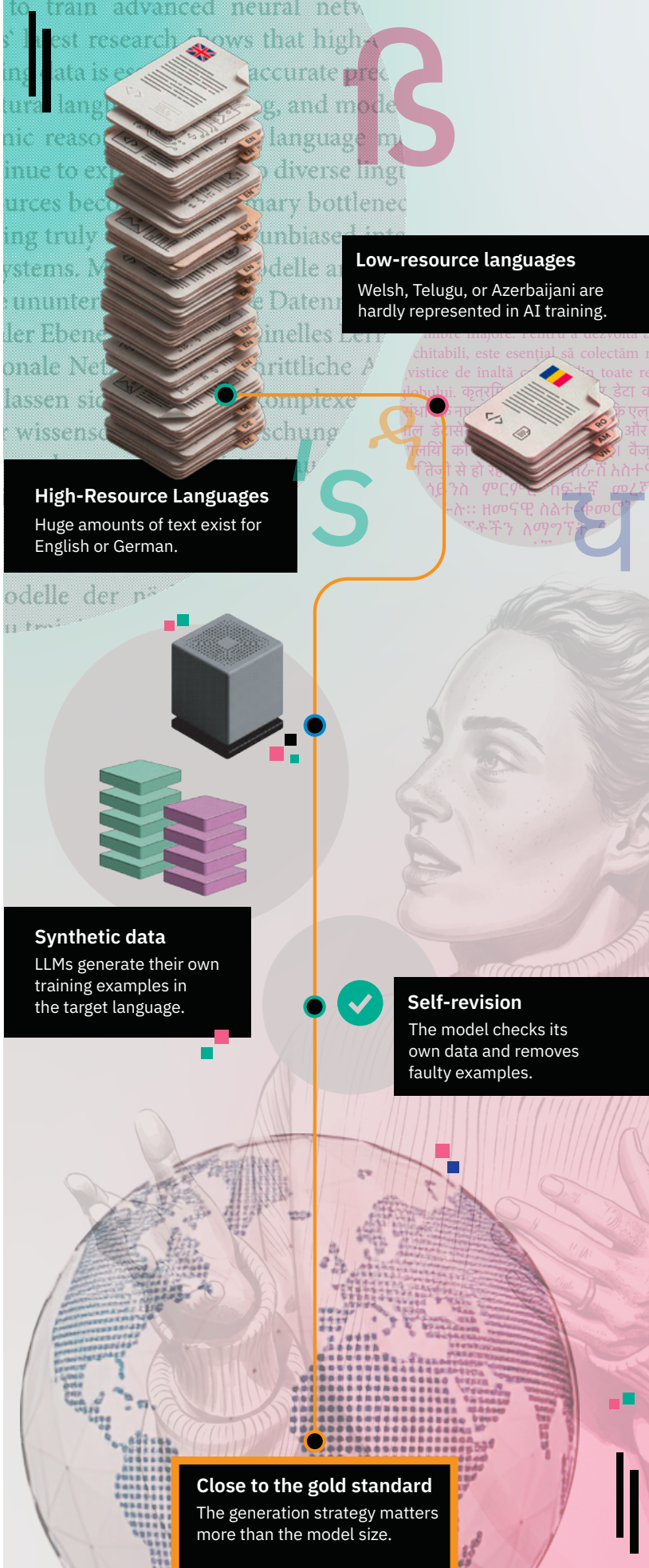
It also became clear that intelligent prompting and revision techniques can partially offset the advantages offered by the large models. This introduces practical possibilities to generate high-quality data even with smaller, more cost-effective open-source models.

« A common assumption is that more parameters automatically yield better results. Our study shows that the approach taken is equally important. With the right prompting and revision techniques, high-quality training data can be generated even for low-resource languages. This is crucial for promoting inclusivity and access to modern AI technology for minority languages. »

Tatiana Anikina, lead author of the study and researcher at the Multilinguality and Language Technology Research Department, DFKI

The results show that the size of a language model is not the sole determinant of the quality of the generated data; a decisive factor is also the generation strategy. The most successful approach used a combination of examples in the respective target language followed by an automated quality check performed by the model itself. The AI reviews its own output and removes erroneous or unsuitable examples.

Despite the promising results, synthetic data does not achieve the same levels of quality across all languages and tasks. Significant performance gaps persist, particularly for severely underrepresented languages. However, the study demonstrates that this gap narrows when data generation is systematically evaluated and subject to appropriate quality control mechanisms.



Papers and authors:

The end of the all-rounder

How AI builds customized high-speed databases

Have you ever wondered why software is often so unwieldy? The reason usually lies in a long-held tenet of software development: „one size fits all.“

Traditional databases like PostgreSQL or DuckDB are like Swiss Army knives – they can handle almost any query. But this flexibility comes at the cost of performance. You pay an invisible “flexibility tax” in the form of processing time and storage consumption for features you don’t even need in your specific case. General-purpose analytical databases are designed for arbitrary workloads and support any query and any database schema. This versatility is both a strength and a weakness: Runtime interpretation, levels of indirection, and abstraction boundaries all introduce additional overhead – even in highly optimized engines.

Realistically, most systems are not asked to execute arbitrary queries and do not require such flexibility. They execute your query: a known set of SQL templates against a known schema, over and over again. A research team at TU Darmstadt (Johannes Wehrstein, Timo Eckmann, Matthias Jasny, and Carsten Binnig) is breaking with this principle. Their revolutionary approach is called “Bespoke OLAP” (Custom Data Analytics) [2603.02001] – and it proposes a radically minimalist future for software development.

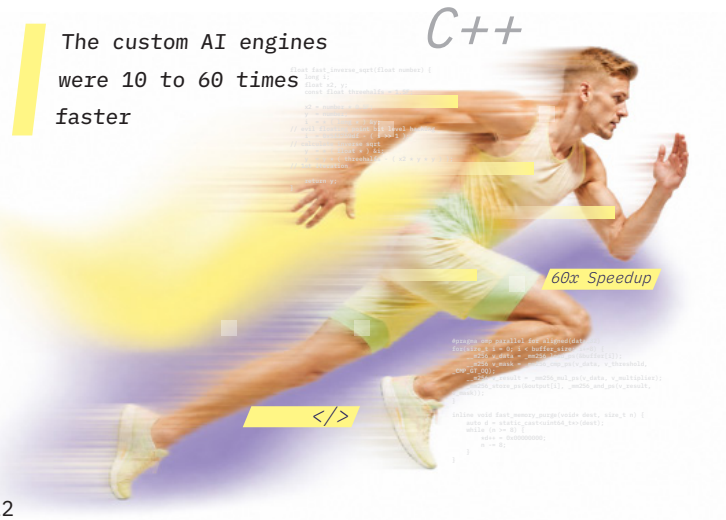
The trick: A custom-made digital solution instead of an off-the-shelf product

Imagine: what if there were no longer any need to install a one-size-fits-all database? Instead, an artificial intelligence system (like Claude or GPT) analyzes the exact data and search queries your company actually uses.

The AI system then constructs an autonomous mini-database in C++ – exclusively for this one task. Since the mini-database does not need to support other types of queries, the overhead of generality is eliminated. Data types, storage layouts, and runtime paths are hard-coded into the database engine.

Up to 60-times faster than the industry competition

The results of initial scientific testing (using the standardized TPC-H benchmark) have sent shockwaves through the tech world: The new AI-generated “One-Size-Fits-One” databases easily outperformed established systems like DuckDB, Clickhouse, or Umbra that have been optimized by human experts for many years. The custom AI engines were 10 to 60 times faster [2603.02001].



Where is the turbo boost seen?

This technology is not a purely theoretical project. "Bespoke OLAP" is a total game changer wherever companies crunch massive volumes of data every night:

E-Commerce & Personalization:

Every night, an e-commerce giant analyzes the click and purchase behavior of millions of customers to personalize the homepage for the following day. Instead of running a standard system at full capacity for hours, an AI system builds a mini-engine in minutes that performs this specific nightly task in a fraction of the time.

Financial Sector & Risk Analysis:

Banks are required to generate regulatory risk and fraud reports on a daily basis. The mathematical queries required for this rarely change. A custom AI database can perform these standard calculations in real time, rather than tying up IT infrastructure for hours.

IoT & Industry 4.0:

In modern factories, thousands of sensors collect machine data every minute for predictive maintenance. A tiny, AI-generated database fits directly onto the on-site control devices and analyzes the data in real time, without having to send it to the cloud first.

What is the current status of the project?

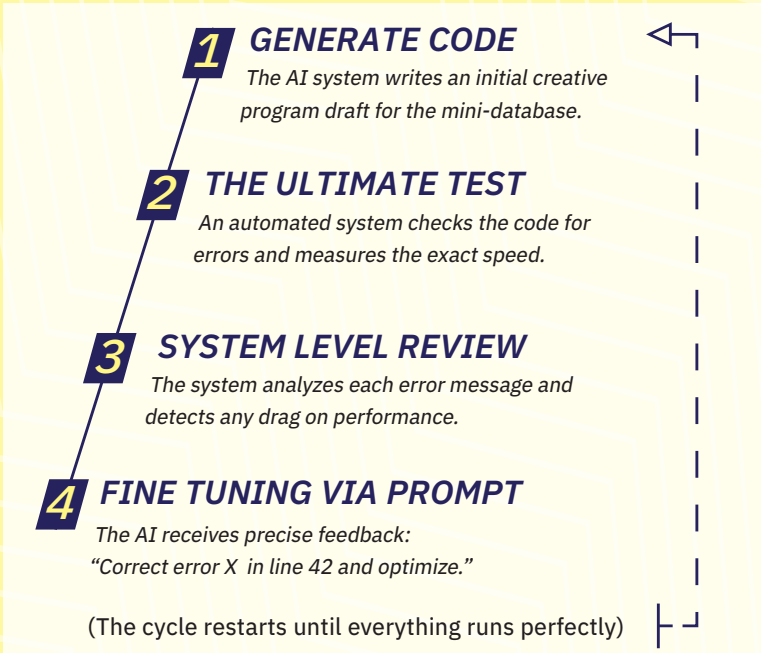
The research is already making waves: following a peer review, the research papers on Bespoke OLAP have been accepted for publication in the IEEE Data Engineering Bulletin (3/2026), one of the most important international journals. It is also accepted for presentation at VLDB (9/2026, Boston), one of the world's leading database conferences. Furthermore, the project was featured in April 2026 on UC Berkeley's renowned Sky Computing blog, invited to the exclusive HPTS Workshop, and discussed in guest lectures at companies including Google and SAP.

Future outlook

"Bespoke OLAP" gives us a glimpse of a new era in software architecture. In the future, AI will no longer merely assist programmers with coding. It will take on the role of chief architect, creating customized digital systems in a matter of seconds, something human developers could previously only dream about.

The highlight:
Continuous testing of AI (The feedback loop)

Historically, AI models fail to generate error-free and extremely fast system code for complex data structures. Researchers solved this problem with an ingenious, closed feedback loop that operates entirely without human intervention.



What's next?

SynnoDB (https://synnodb.com/) is the spin-off company where Prof. Carsten Binnig's team aims to exploit the full potential of the technology and bring the innovative Bespoke-OLAP technology to market. Conventional databases are like all-purpose vehicles. SynnoDB constructs a custom-built racing car for each customer, one precisely tuned to the data and queries required, while delivering significantly faster performance. The company's website currently features a "try-it-yourself" demo that allows users to create synthetic mini-databases, as well as a live playground where you can:

- Run queries on a real engine generated by SynnoDB.
- Compare performance against DuckDB and Umbra.
- View the generated C++-source code.
- See how the engine was built.

Papers and authors:

X-Hacking

When explainability is misleading

DFKI researchers reveal how explainable AI can become a methodological trap: When AutoML selects – from among many equally good models – the very variant that provides an explanation that perfectly fits a desired narrative, transparency itself is suspect.

XAI methods are intended to make machine learning systems more transparent. However, precisely because models provide plausible explanations for their decisions, a problem arises – described in our paper as "X-Hacking." The problem is that explanations can be deliberately steered in a desired direction even though the models perform almost equally in terms of predictive accuracy.

The starting point is a simple yet far-reaching observation: multiple models achieve similar accuracy on a given dataset but offer very different justifications for their results. Anyone choosing from this set of "plausible" models can favor precisely the explanation that supports a desired narrative. Our paper extends the principle of p-hacking – well-known in statistics as a way to get the most pleasing results – to explainable AI: It optimizes not only the prediction but also the justification for the result.

The problem becomes particularly relevant with automated search and optimization methods. AutoML greatly facilitates model development, while also expanding the range of possible decisions. This is the scientific crux of X-Hacking: When many models, preprocessing steps, and

ICML: Where machine learning meets applications

The International Conference on Machine Learning (ICML) is one of the top three leading conferences in machine learning (next to NeurIPS and ICLR). In scope, it is characterized by theoretical analysis, algorithmic innovation, and statistical learning – with strong interfaces to reinforcement learning, robotics, and optimization theory. The conference provides a central hub for researchers in AI, statistics and data science as well as application-oriented fields like computer vision, computational biology, speech and robotics. At ICML 2025, the DFKI team led by Prof. Sebastian Vollmer presented a paper with the title "X-Hacking: The Threat of Misguided AutoML." It applies the principle of p-hacking to explainable AI and shows how AutoML searches pipelines systematically for models with desired, but potentially misleading explanations.

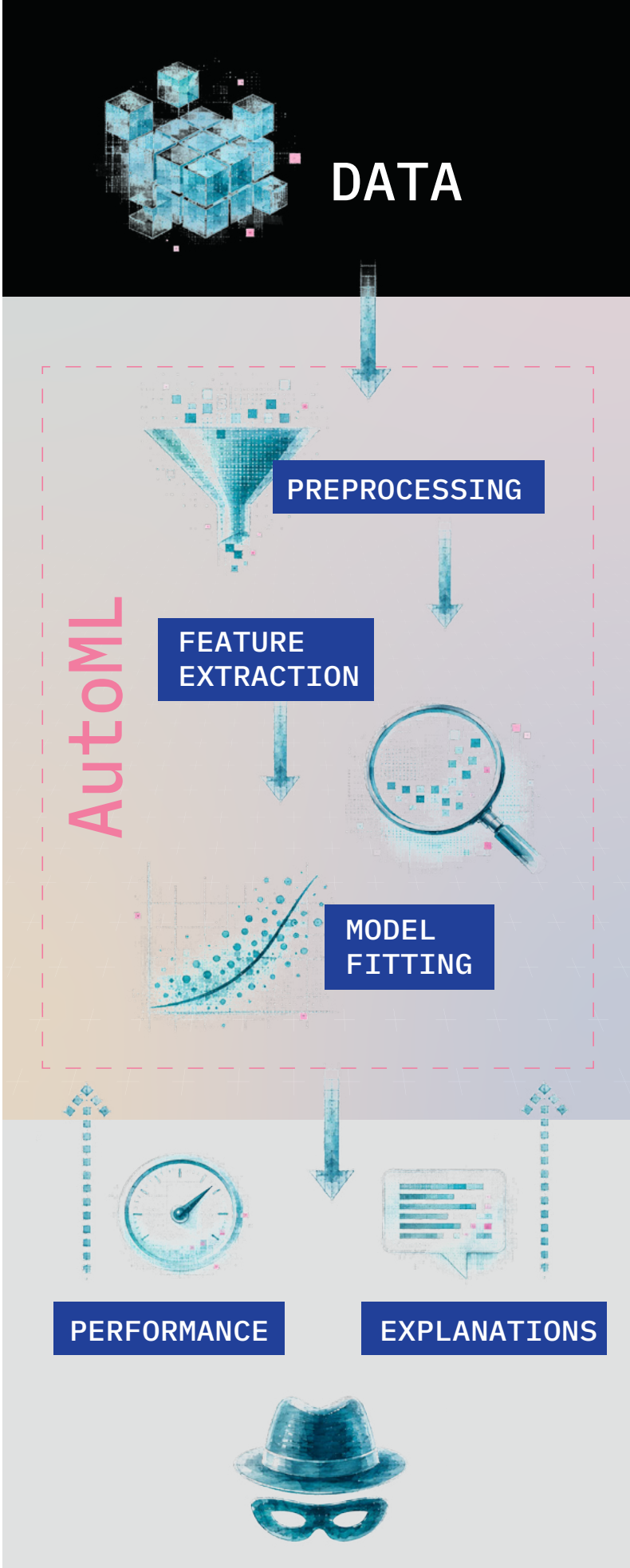
« At a time when AI explains decisions, but does not always understand them, we as scientists must take responsibility for the depth of the explanations – and recognize their limitations. The goal is a scientific culture that prioritizes not only accuracy, but also honesty in explainable AI. »

Prof. Dr. Sebastian Vollmer,
Head of Data Science and its Applications Research Department at DFKI

parameter settings are available, the search for a "suitable" explanation can become a part of the systematic analysis process. Explainability is no longer automatically seen as a corrective measure, but rather as something that must itself be methodologically tested.

This takes on considerable importance for sensitive fields of application, like medicine, social research, and the justice system, where explainable models are often considered a prerequisite for trust, transparency, and responsible decision-making. However, if a plausible explanation is selected from a range generated by models of similar high quality, a reference to interpretability is no longer adequate. This point of tension is where the Trustworthy AI debate sets in, and our contention is that what matters most is not only the accuracy of the model but also the integrity of the entire process by which explanations are generated.

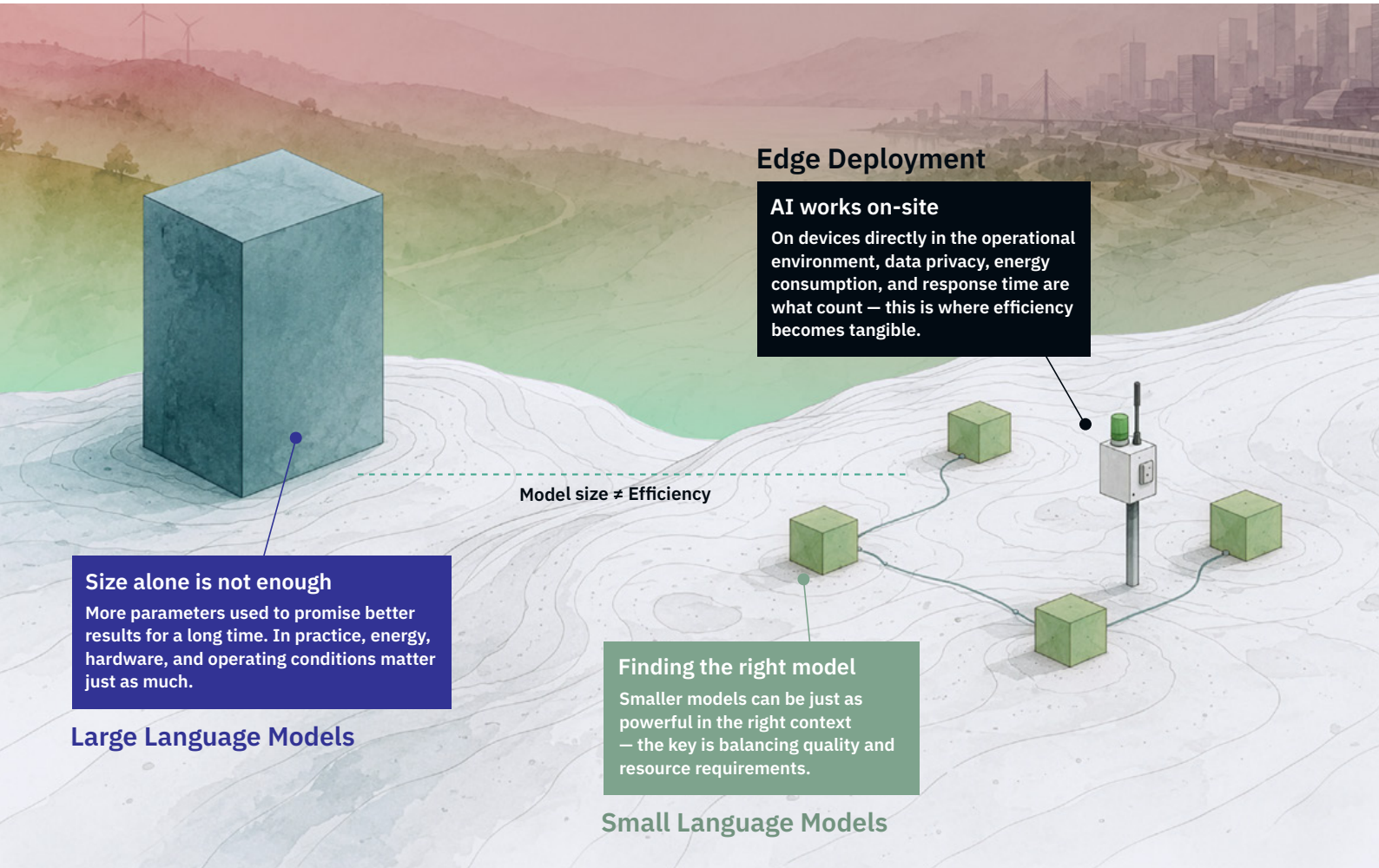
The paper does not conclude with a simple problem description. It advocates for a more reflective scientific application: explanatory histograms that visualize the range of possible explanations, complete documentation of the pipeline, and greater awareness that automated model selection not only enhances efficiency but can also introduce new biases. X-hacking is not some sideshow of explainable AI, but a fundamental question of scientific integrity regarding the use of automated model search.



Efficiency is the new benchmark for language models

Why size alone is not the crucial factor when choosing a language model

The development of large language models has long followed a seemingly simple logic: more parameters, more computing power, and larger models mean better results. This paradigm of scale is what shaped AI research in past years. However, as more language models find their way into business and social applications, it becomes clear that size is not always better and practical utility matters more.



« We are creating the scientific basis for the efficient use of language models. We will present our findings, as reported in the paper titled “Mapping the Efficiency Landscape of Small Language Models”, at IJCAI 2026. We demonstrate that it is not the size of the model that is critical, but rather how efficiently the model meets the specific requirements of the current application. »

Fabian Reichwald, Research Department Intelligent Markets at DFKI

Why larger is not automatically better

Whether in industrial processes, on edge devices, or in sensitive data environments: language models must perform under real-world conditions and requirements. For example, the largest available model is not always the best choice for a medium-sized enterprise – such as one summarizing technical documents or automating the processing of customer inquiries. In this case, the model that performs the task on existing hardware while striking the best balance between output quality, energy consumption, and response time may be a better choice.

This marks a shift in the scientific focus. Instead of evaluating models solely on the basis of size or performance, the focus is now on practical considerations. How much computing power does a task require? When is higher energy consumption justified by a measurable gain in quality, and when does it merely drive up costs and resource consumption?

More than performance

Efficiency is no longer a secondary consideration, but a scientific variable in its own right. For a company that intends to run a language model locally, simply looking at the model’s size is not enough. What matters most is which model solves the specific task with the fewest possible resources. Efficiency measurements rely on the relationship between performance, energy consumption, and application context.

Under this scrutiny, many common assumptions prove to be invalid. Smaller models do not automatically operate more efficiently than larger ones. Additional parameters or higher energy consumption do not necessarily lead to better results. Well-informed decisions require a systematic consideration of all influencing factors.

Transparency of well-informed decisions

The next step in the development of language models is not simply to build more powerful models. Equally important are the scientifically grounded criteria on which to base the selection of the right model for the specific task.

A model’s practical value is determined not by its size alone, but by its suitability for the specific application context. Efficiency is becoming a key criterion when selecting language models, particularly where technical, economic, and environmental requirements converge.

Bridging Generative AI and Robotics

From task understanding to reliable actions

How does data translate into reliable robot actions in the physical world? To achieve this, a new approach introduces an additional architectural layer between generative AI and robotics. Instead of translating language or images directly into robot movements, large AI models first generate structured task representations that are supplemented with physical requirements. This creates a link between task interpretation and robust robotic actions.

Language models can understand complex relationships and generate plausible responses. Robotic systems must also contend with physical forces, dynamics, and uncertainties. The new architectural layer makes AI-generated task representations practical for robotic task execution in real-world scenarios. The approach also gives more flexibility to highly dynamic robotic systems to adapt to different environments.

The new approach is based on a new form of task representation. Instead of directly translating an instruction such as „move an object to a specific location“ into control commands, AI first analyzes the task. It captures the objective, relevant conditions, and the characteristics of both the robot and its environment, and then uses this information to generate a semantic task representation.

Image AI-generated based on a photograph.

The Unitree G1 is one of the research platforms used in the ActGPT project.

« The real challenge is not getting robots to understand instructions, but in ensuring they can execute them reliably and robustly within existing physical constraints. Generative AI helps to understand the context of natural language and formulate the problem, while physics determines whether the solution actually works. »

Shubham Vyas,
DFKI Robotics Innovation Center

This representation is subsequently converted into a machine-readable format by a specially developed domain-specific language (DSL). The DSL serves as an interface between generative AI and the mathematical optimization methods used in robotics. Once provided with the movement requirements, an optimal control problem (OCP) is derived. This problem is then solved using established control methods to generate appropriate movements for the task, while ensuring that it is also physically possible.

This approach represents a broader trend in AI research: What matters most now is not simply how machines understand and model their environment, but how reliably they translate that knowledge into real-world actions.

Research project ActGPT

The "ActGPT" research project at the DFKI Robotics Innovation Center is funded by the Federal Ministry of Research, Technology, and Space (BMFTR). Researchers are making AI-generated task representations useful for real-world application scenarios by developing methods to integrate generative AI with robust robotic systems. The aim is to support the reliable execution of complex robotic tasks and to simplify the development of control strategies.

The ActGPT Workflow

1 Input

- Language
- Images
- Sensor data

2 Generative AI

Understands the task

New Architectural Layer

Domain-Specific Language (DSL)

Describes:

- Goals
- Cost functions
- Constraints

Optimal Control Problem (OCP)

3 Optimal Control

Solves the OCP
Computes optimal trajectories

4 Robot

Executes the motion

ACT GPT

The new architectural layer. AI models translate tasks into structured physical problems, directly connecting semantic knowledge with robust robotics.

Image AI-generated based on a photograph.

Shaping the future of AI-based robotics through collaboration

DFKI combines scientific excellence, system competence, and real-world testing

The next stage in the development of artificial intelligence is playing out in the real world – with robots capable of understanding their surroundings, learning from experience, and acting safely. The marriage of AI and physical action is one of the main scientific challenges in AI research today and is being pursued under the term "Embodied AI."

DFKI is making a significant contribution to the advancement of AI-based robotics in Germany through its work at the intersection of artificial intelligence, autonomous systems, and robotics. Its strengths are in basic scientific research, technological systems expertise, and real-world testing.

From AI methods to intelligent robot systems

At various research departments and locations, DFKI investigates key issues in intelligent robotics – from perception and learning, through planning and autonomous decision-making processes, to human-robot interaction and practical applications.

The work concerns individual AI methods and technological components as well as the development and testing of complete robotic systems. The coordinated interaction of different technologies is crucial: software, sensors, controls, and physical implementation must all come together in a functioning overall system. Only then can autonomous robots behave safely and flexibly in dynamic surroundings.

Holistic approach and system competence

This approach is particularly evident at the Robotics Innovation Center in Bremen, where complex robotic systems are developed and tested under realistic conditions.

The research platforms developed here include: autonomous underwater robot systems and space rovers, walking robots and humanoid systems, robotic exoskeletons, and individual components like manipulators and grippers. This facilitates the study of key challenges in intelligent robot design – from autonomous navigation systems to safe interaction with humans.

The importance of research platforms and test facilities

A key feature of robotics research is the research platform. In-house systems make it possible to study new concepts and technologies at the system level and to investigate the interaction of various components. Commercial platforms improve the research results by enabling reproducible experiments and comparisons under standardized conditions.

« Robotics is a field that requires the convergence of many disciplines. We address every level of a robotic system: from hardware and electronics to software development and innovative AI methods. Few institutions offer such a broad range of expertise. »

Dr. Sirko Straube, Deputy Head,
DFKI Robotics Innovation Center

Equally important are specialized testing environments where new methods and systems can be tested under realistic conditions. With the Maritime Exploration Hall and the Multifunctional Hall in Bremen – which, among other features, include an artificial lunar crater landscape – and the SmartFactory in Kaiserslautern, the DFKI has unique infrastructure for various areas of robotics research. These facilities enable the systematic validation of new approaches and the evaluation of their performance in realistic test scenarios.

Advancing Robotics Research Together

DFKI is actively involved in national and international research networks. The continued development of intelligent robotics demands such collaboration – across various scientific disciplines and technological expertise and beyond the boundaries of individual institutes and fields of study.

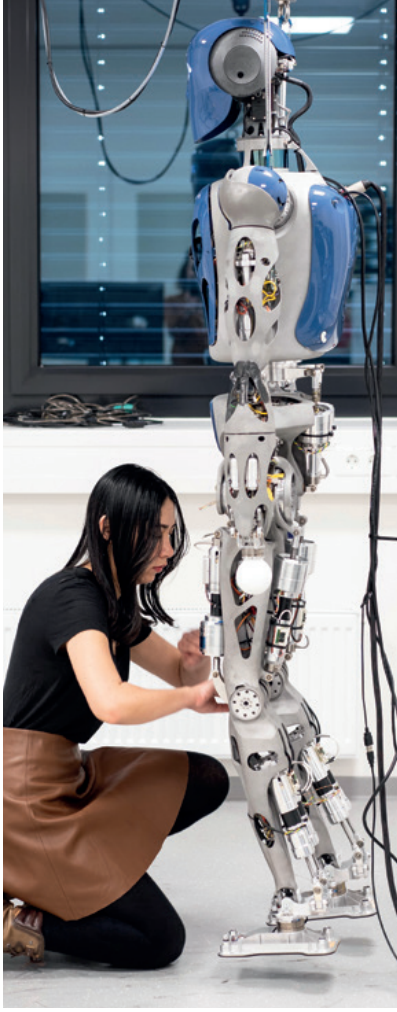
The Robotics Institute Germany (RIG) presents a prime example of this collaboration. DFKI offers its expertise in autonomous robotics and artificial intelligence as well as its experience in the development and testing of complex systems. Networking with leading robotics hubs across Germany pools diverse competencies and creates a framework for joint research activities.

Through the combination of scientific excellence, technological system expertise, and interdisciplinary collaboration, DFKI is able to make significant contributions to the study of key issues in AI-based robotics. In this way, DFKI is helping to shape the next generation of intelligent robotic systems.

Image AI-generated based on a photograph.

DFKI is a partner in the Robotics Institute Germany (RIG)

The Robotics Institute Germany (RIG) is an initiative of the Federal Ministry for Research, Technology, and Space (BMFTR) for the purpose of networking the leading robotics hubs in Germany. The aim is to pursue scientific excellence, increase international visibility, attract talent, and accelerate technology transfer in AI-powered robotics. DFKI participates in the RIG through three research departments: the Robotics Innovation Center in Bremen and the departments of Intelligent Networks and Innovative Factory Systems in Kaiserslautern. DFKI's contributions include the provision and use of research facilities, training and professional development opportunities for specialists and managers, participation in international robotics competitions and challenges, and networking with industry partners.



DFKI @ IJCAI-ECAI 2026

Perspectives on the AI of Tomorrow

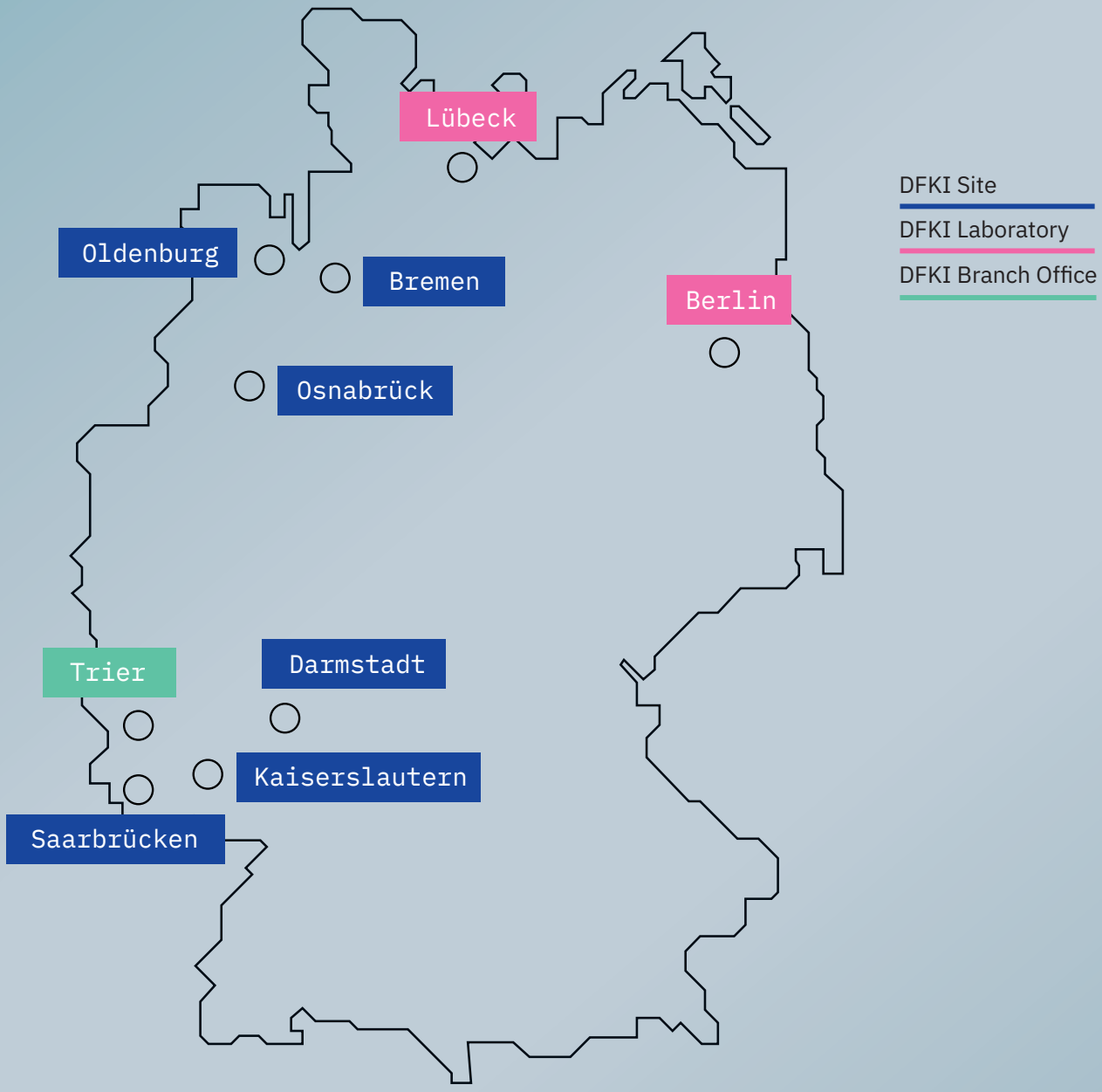
DFKI is well represented at one of the world's most respected conferences on Artificial Intelligence. The International Joint Conference on Artificial Intelligence (IJCAI) 2026 has accepted twelve papers published by the German research center. The papers reflect the scientific diversity of the institute – ranging from foundation models and computer vision to robotics, semantic technologies, and explainable AI as well as medical applications, environmental monitoring, and software engineering.

Together, these papers highlight the topics currently shaping international AI research. How can we design intelligent systems to be more efficient, more transparent, and more interactive? How can we help them to better understand complex relationships and reliably support people across a wide range of application scenarios? These DFKI papers provide insights into the next generation of intelligent AI systems.

For more information, please scan the QR code

www.dfki.de

AI-generated cityscape.



Imprint

dfki ai next; Issue August | 2026; **Published by:** Deutsches Forschungszentrum für Künstliche Intelligenz GmbH (DFKI); Trippstadter Str. 122, D-67663 Kaiserslautern; **Phone:** +49 631 20575 0; **E-Mail:** news@dfki.de; **Editors:** Jeremy Gob (responsible), Heike Leonhard, Udo Urban, Andrea Fink, Maria Santos; **Editorial review:** Armindo Ribeiro; **Layout:** Lando Lehmann, Juan Mejia; **Production:** One Vision Design, Saarbrücken; **Cover:** AI-generated, graphic design by DFKI.

[AI] Image contains AI-generated visual elements. Concept, compositing, and design: **DFKI**. For photorealistic images, the label is also placed directly on the image.

Credits

Title: DFKI, Lando M. Lehmann [AI]; **Page 5:** DFKI, Oliver Dietze; **Page 6:** DFKI, Lando M. Lehmann [AI]; **Page 8:** DFKI, Juan D. Mejía [AI] (Person); Louisa Howard, Public Domain (Microscopy); **Page 9:** DFKI, Juan D. Mejía [AI]; **Page 11:** DFKI, Lando M. Lehmann [AI]; **Page 12:** DFKI, Juan D. Mejía [AI]; **Page 14:** DFKI [AI]; **Page 15:** DFKI, Lando M. Lehmann [AI]; **Page 16:** DFKI, Lando M. Lehmann [AI]; **Page 18:** DFKI, Juan D. Mejía [AI]; **Page 19:** DFKI, Juan D. Mejía [AI]; **Page 21:** DFKI, Juan D. Mejía [AI]; **Page 22, top:** DFKI, Thomas Frank; DFKI, Annemarie Popp; **Page 22, bottom:** DFKI [AI].