relAI

Building trustworthy AI for a safer tomorrow

relAI [rɪ'laɪ]
Konrad Zuse School
of Excellence in Reliable AI
2024–2025

The Konrad Zuse School of Excellence in Reliable AI is a vibrant community, composed of **top talents from around the world and Fellows who are the most renowned researchers in their field.** Together, they have taken on one of the most important tasks of our time: **making Artificial Intelligence safe and reliable.**

Directors

Dear readers,

Three years into the journey, our Konrad Zuse School of Excellence in Reliable AI (relAI) has taken significant steps towards our goal to become the international beacon we want to be in the field of reliable artificial intelligence.

Our student body has grown not just in size, but also in diversity and expertise. We celebrated excellent research results and prizes, we sent our students out into the world to exchange with outstanding international researchers at conferences and at top universities, we introduced a new research area as well as new curricular activities, and we continued to foster a vibrant community through numerous events.

And we have continued to grow our relAI family not just in terms of students: More scientists from the two German top ranked universities Technische Universität München (TUM) and Ludwig-Maximilians- Universität München (LMU) have become relAI Fellows to support this joint school, and more companies have become industry partners to facilitate transfer and exchange.

Our 2025 report will provide a comprehensive overview of our achievements over the past year and introduce you to all the key developments.

We would like to once again express our sincere gratitude to the German government, represented by the German Academic Exchange Service (DAAD) and the Bundesministerium für Forschung, Technologie und Raumfahrt (BMFTR), for funding our Konrad Zuse School and for enabling this visionary endeavor. Our heartfelt thanks also go to the Presidents of TUM and LMU, Prof. Dr. Thomas Hofmann and Prof. Dr. Dr. h.c. Matthias Tschöp, for their ongoing support and collaboration. Last but not least, we extend our warm thanks to everyone contributing to the success of relAI—in every role and capacity. Your dedication makes all the difference.

Sincerely,

Prof. Dr. Stephan Günnemann (TUM)
Prof. Dr. Gitta Kutyniok (LMU)
Directors of the Konrad Zuse School in Reliable AI

Shaping the **future** of reliable **AI**

The ongoing technological revolution is profoundly shaped by significant advancements in **artificial intelligence** (AI).

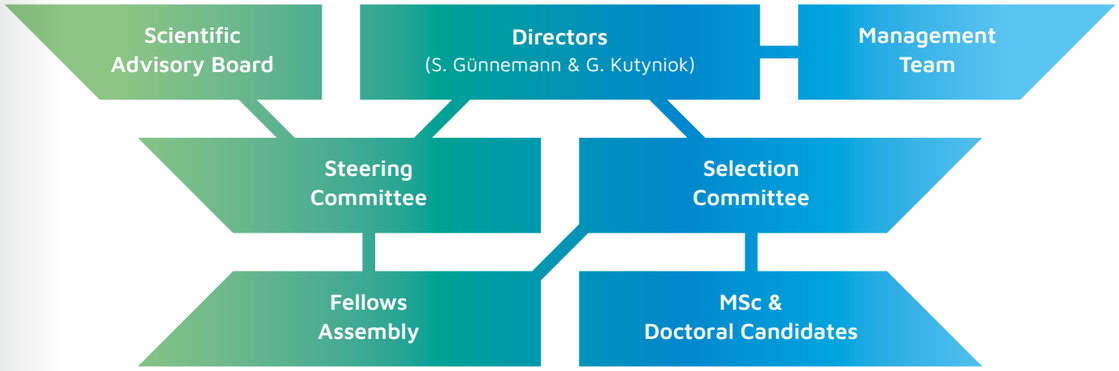
While the transformative potential of AI is widely acknowledged, persistent concerns regarding its **reliability** continue to pose a major obstacle to its widespread adoption across industry and society.

At the Konrad Zuse School of Excellence in Reliable AI (relAI), we understand that ensuring the safety, security, and privacy of AI systems is imperative—particularly in areas of public interest, where the stakes include safeguarding human life and protecting sensitive data. Hence, our mission is to train future AI experts who uniquely blend technical proficiency with a deep understanding of AI reliability. Through an innovative and interdisciplinary program, relAI provides top international candidates with advanced scientific training, business acumen, and practical industrial experience, preparing them for roles in both industry and academia. Additionally, we conduct cutting-edge research to ready AI for deployment in critical applications.

In the fourth year of its existence, relAI has grown to its full size. We have welcomed new cohorts of excellent Master and PhD students from all over the world, and we have sent off the first alumni. In this report, we will show relAI's developments over the past year, introduce the great work that's being done here and let some of the members of our relAI family speak for themselves.

How we are organized

The Konrad Zuse School of Excellence in Reliable AI is built on a robust and collaborative governance structure. This framework ensures effective oversight and strategic direction for all our initiatives.



At the helm are our **Directors**, Prof. Dr. Stephan Günnemann (Technische Universität München, TUM) and Prof. Dr. Gitta Kutyniok (Ludwig-Maximilians-Universität München, LMU) who provide leadership and vision. They are supported in day-to-day operations by the **Management Team**, which handles all administrative functions and coordinates activities across the school.

Strategic direction comes from our **Steering Committee**, a key body comprising the directors and senior representatives from our four research areas at both TUM and LMU. This committee ensures our activities are aligned with the strategic goals.

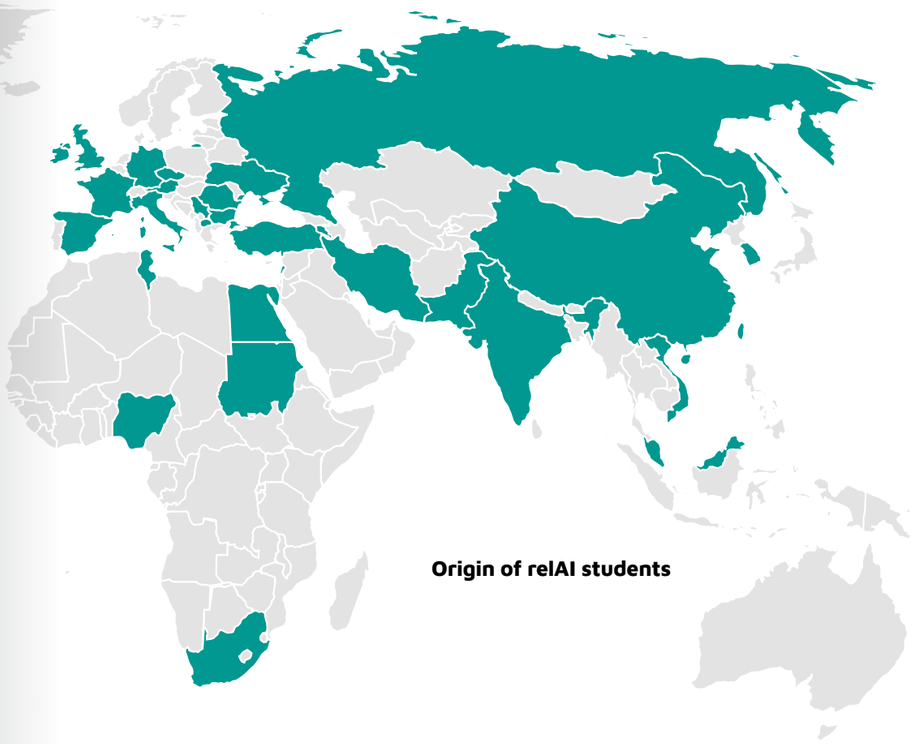
The **Selection Committee**, which includes all Steering Committee members plus four additional Fellows from both universities, is responsible for the vital task of recruiting and selecting top talents for our Doctoral positions and MSc scholarships.

To maintain the highest scientific standards, we rely on the guidance of our **Scientific Advisory Board**. This board is composed of distinguished international scientists who offer expert insights to ensure our research and development priorities remain at the leading edge.

Finally, the **Fellows Assembly** brings all relAI Fellows together annually to review progress, discuss future directions, and elect new committee members, ensuring a transparent and participatory governance model. This collaborative spirit extends to our **students**, who take an active role in the school's community by managing the relAI Blog, organizing the alumni network and extending the relAI Wiki, and planning various excursions and talks.

Our relAI Family

Being dedicated to advancing AI research and innovation, relAI achieves its vision of creating reliable solutions through the distinct perspectives and shared commitment of our entire community, which includes MSc and Doctoral researchers, Fellows, and our industry and academic partners.



Origin of relAI students

Students

Our MSc and Doctoral students are the very heart of our school. Not only do they drive research in reliable AI, but they also enrich the entire community by fostering stimulating discussions, promoting our work through outreach activities, and expanding the relAI network via internships and academic exchanges.

Since 2022, our community has grown significantly, with 42 MSc and 64 Doctoral students joining us through four application calls. Hailing from 35 countries across the globe, our students bring a truly international perspective to our work.

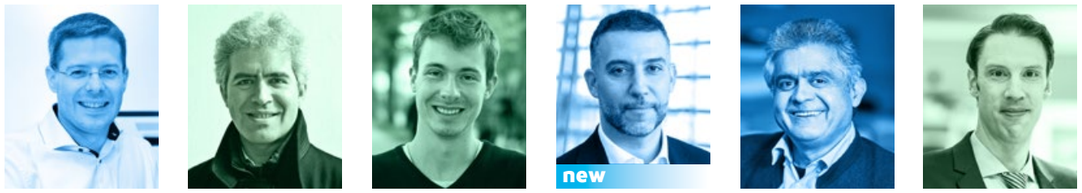
The high standard of their research is already evident: our students have contributed to over **160 publications to date**. The strong academic output highlights the quality of their work. This year, the first relAI PhD candidates will complete their degrees, and we are proud that many of our MSc graduates are continuing their academic journeys with us or pursuing doctoral studies at other renowned international universities. A small number have already embarked on successful careers in industry, reflecting the diverse pathways our program enables them to take.

1

2

3

4



5

1 Albu-Schäffer, Alin TUM
Althoff, Matthias TUM
Bauer, Stefan TUM, Helmholtz AI
Bhatotia, Pramod TUM
Bischl, Bernd LMU
Butz, Andreas LMU
Buyx, Alena TUM
Cremers, Daniel TUM

2 Drton, Mathias TUM
Eckert, Claudia TUM
Feuerriegel, Stefan LMU
Fortuin, Vincent TUM, Helmholtz AI
Ghoshdastidar, Debarghya TUM
Günnemann, Stephan TUM
Heckel, Reinhard TUM
Hirche, Sandra TUM
Hüllermeier, Eyke LMU
Ingrisch, Michael LMU

3 Kaissis, Georg TUM
Kasneci, Enkelejda TUM
Kasneci, Gjergji TUM
Kauermann, Göran LMU
Kern, Christoph LMU
Kilbertus, Niki TUM, Helmholtz AI
König, Alexander TUM
Kreuter, Frauke LMU
Kuhn, Jochen LMU
Kutyniok, Gitta LMU

Fellows

Our Fellows are the academic architects of the relAI program. As distinguished experts in their fields, they provide the strategic vision and intellectual framework for our mission to advance research and innovation as well as create reliable and impactful AI solutions. Their expertise guides our research across five primary research areas: Mathematical and Algorithmic Foundations of Reliable AI, Medicine & Healthcare, Robotics & Interacting Systems, Algorithmic Decision-Making, and Learning & Instruction. Through their deep knowledge and dedicated supervision, the Fellows are instrumental in providing academic content, mentoring our students and shaping the development of the next generation of AI leaders.

4 Lasser, Tobias TUM
List, Christian LMU
Maly, Johannes LMU
Masia, Lorenzo TUM
Navab, Nassir TUM
Ommer, Björn LMU
Rückert, Daniel TUM
Rügamer, David LMU
Schmidt, Albrecht LMU
Schnabel, Julia TUM, Helmholtz Munich

5 Schöllig, Angela TUM
Schuller, Björn TUM
Schütze, Hinrich LMU
Sterkenburg, Tom LMU
Theis, Fabian TUM
Tresp, Volker LMU
Vieluf, Solveig LMU
Wachinger, Christian TUM
Wiestler, Benedikt TUM
Zöller, Mark LMU

Academic Partners

relAI cooperates with a number of academic Institutions all over the world. These collaborative relationships are intended to enhance the quality, scope and impact of our research endeavors, facilitate networking for our students and Fellows, and allow for knowledge exchange and research stays.

relAI has supported 14 research visits for students to prestigious international research centers, including Harvard University, the University of California-Berkeley, and the University of Pennsylvania. This support has been instrumental in fostering research collaboration, resulting in significant achievements, such as an Outstanding Paper Award at ICML 2025.

Later in this report, we will let some of our students speak about their research stays and share their perspectives on the benefits they gained.

Data Science at Uni Vienna

Center for Intelligent Systems (CIS)
at École Polytechnique Fédérale de Lausanne (EPFL)

Center for Responsible AI Technologies
coordinated by Munich School of Philosophy

Helmholtz München

Alan Turing Institute

Center for AI Research (CAIRE)
at Hong Kong University of Science and Technology (HKUST)

NSF AI Institute for Advance in Optimization
at Georgia Institute of Technology (AI4OPT)

Center for Data Science (CDS)
at New York University (NYU)

Center of Data Science and AI Research (CeDAR) at University of California, Davis (UC Davis)

Center for Statistics and Machine Learning (CSML) at Princeton University

Mathematical Institute for Data Science (MINDS) at John Hopkins University

Oden Institute for Computational Engineering and Sciences
at University of Texas at Austin

Rhodes Information Initiative (RII)
at Duke University

Stanford Data Science Centre (SDS)
at Stanford University

Industry Partners

Cooperation with industry partners is another important pillar supporting the structure of our school. Part of relAI's curriculum includes internships for MSc and Doctoral students. Doctoral students have one industry member on their Transdisciplinary Thesis Advisory Committee (TTAC), so they receive feedback on their research and personal development from a transferability point of view. Many renowned companies have already partnered with us, and the network is still growing by two to three new partners per year.

new **Airbus**

Allianz

BMW

Bosch

Celonis

Denso

Fraunhofer Institute for
... Applied and Integrated Security AISEC
... Cognitive Systems IKS
... Integrated Circuits IIS

new **GE HealthCare**

Google

Imfusion

Infineon

Linde

Munich Re

new **QuantCo**

SAP

Siemens

Siemens Healthineers

thyssenkrupp

UnternehmerTUM

Volkswagen

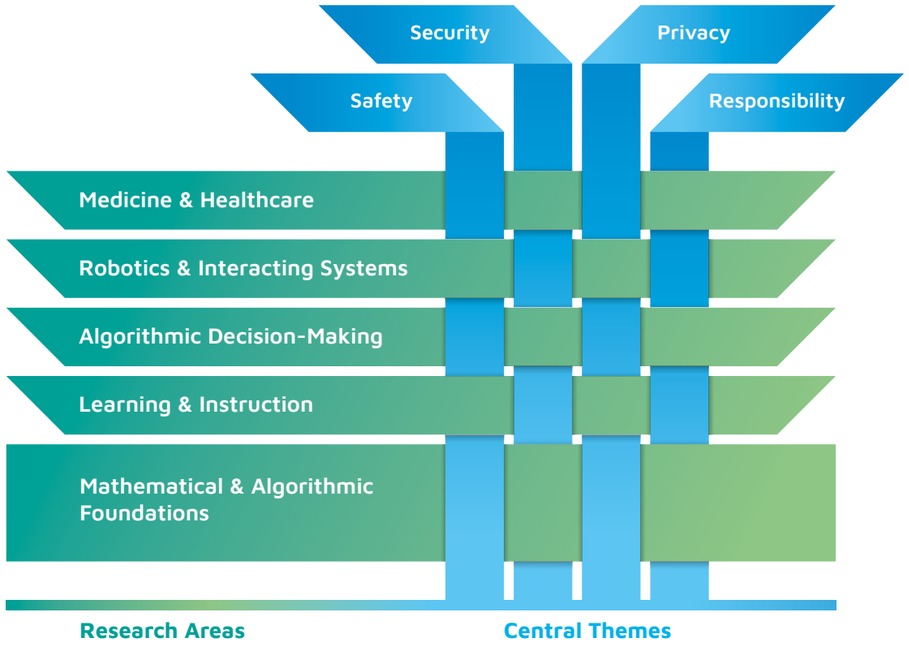
The pursuit of **Reliable AI**

At its core, **relAI's mission** is to train the next generation of AI experts in Germany. We are educating professionals who possess not only technical brilliance, but also a deep understanding of the critical importance of AI's trustworthiness.

As a central component within the Munich AI ecosystem, relAI is committed to advancing artificial intelligence that is fundamentally **safe, secure, and responsible**, with a strong focus on protecting individual **privacy**. We champion the responsible use of AI to exploit its immense potential for the benefit of humanity, to foster crucial discourse on the ethical and societal implications of machine learning, and to reinforce the strength of "AI made in Germany."

Working Together to Create Impact

Our research projects are powered by a collaborative effort, bringing together professors, post-docs, Doctoral candidates, and MSc students across various academic departments and research centers. The impact of our work extends beyond academia; we actively disseminate our findings to key innovative industries, including medicine and healthcare, robotics and interactive systems, and algorithmic decision-making.



The relAI scientific program drives the **end-to-end development of reliable AI**, covering different branches of **applied research** on the basis of profound **mathematical and algorithmic foundations**. This theoretical grounding is a key differentiator for relAI. We believe that true reliability requires a rigorous formal description of properties and **provable guarantees**. It is these guarantees that will build the trust and confidence necessary for the widespread and practical adoption of AI, without reservations.

Our research program integrates the Mathematical and Algorithmic Foundations of reliable AI with deep domain knowledge across four core application areas (as visualized in the figure on the left): Medicine & Healthcare, Robotics & Interacting Systems, Algorithmic Decision-Making, and Learning & Instruction. These domains are of significant importance to Germany and represent areas where the need for reliable AI methods is most pressing. To create a revolutionary impact in education, where reliability is essential, relAI has recently introduced the field of "Learning & Instruction." For more details on this new area, we invite the reader to explore the following section. By focusing on these high-impact topics, our school's research directly addresses key societal demands and public interest.

Our school's research is driven by a core mission: to build AI that is trustworthy and beneficial to society. To achieve this, each of our research areas addresses four key themes that constitute the central challenges of reliable AI:

Safety We ensure that AI systems, such as robots, do not put people in danger or cause any harm.

Security We focus on making AI resilient against external attacks, threats, and information leaks. For example, we prevent manipulation of decision-making systems for criminal purposes.

Privacy We develop methods to protect the confidentiality of sensitive (individual) data and information, especially for applications like medical AI systems that handle patient data.

Responsibility We work to ensure AI systems align with societal norms and ethical principles while taking human needs into consideration, for example, by explaining AI decisions and protecting individuals against discrimination.

By addressing these core themes and combining foundational research with practical applications, we foster an even stronger interdisciplinary environment at relAI.

Since its inception, relAI has published findings in over 180 scientific publications. To showcase the breadth and depth of our work, the following section highlights a representative paper for each area.

Mathematical & Algorithmic Foundations

Achieving truly reliable AI requires a profound understanding of its underlying mathematical and algorithmic foundations. A primary challenge we address is bridging the persistent gap between AI theory and practice—a key obstacle to creating the robust, comprehensive guarantees needed for critical applications. To support our mission of reliable AI, our research tackles two primary challenges.

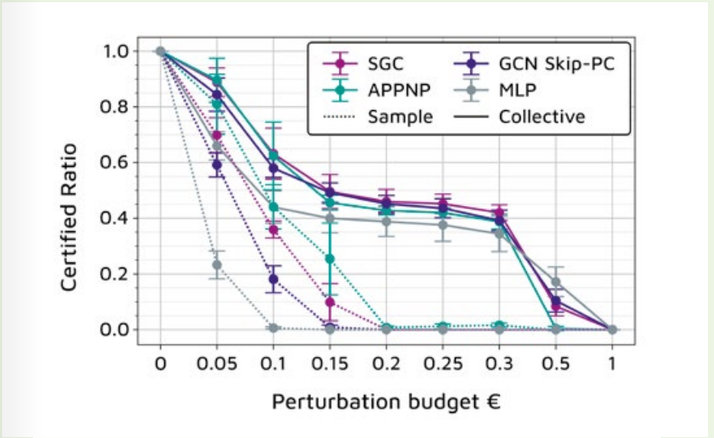
First, we are dedicated to establishing **theoretical guarantees** for AI. Our work here includes expressivity of AI models, analysis of learning algorithms, generalization capabilities of trained AI systems, and evaluating the generalization and robustness of trained systems. A key focus is on deriving concrete error bounds and certifications, which is especially challenging for complex, novel architectures like graph neural networks and transformers.

Second, we are advancing the **algorithmic foundations** of AI to directly support reliability in real-world contexts. Our research spans critical topics such as IT security, federated learning, distributed systems, and causal modeling, ensuring a tight link between our theoretical work and its practical application across all our domains.

These theoretical guarantees and algorithmic foundations are explored in the paper "Exact Certification of (Graph) Neural Networks Against Label Poisoning," which is the result of a cooperation of multiple relAI Fellows and Doctoral students. It was selected as a Spotlight Paper at the International Conference on Learning Representations (ICLR 2025) and as an Oral Presentation for the VerifAI Workshop.

Sabanayagam, M., Gosch, L., Günnemann, S., Ghoshdastidar, D. (2025). » Exact Certification of (Graph) Neural Networks Against Label Poisoning. In "The Thirteenth International Conference on Learning Representations".

Neural networks, including those used for classifying structured data (e.g., Graph Neural Networks or in short GNNs), are highly susceptible to label poisoning. This is a form of data corruption, where a small number of training labels are corrupted to significantly degrade the model's performance—a process that can happen naturally through noise or erroneous data, or be deliberately induced by malicious actors. The challenge lies in creating a defense that can provide provable guarantees, or "certificates," that the model's predictions will remain unaffected even in worst-case corruption scenarios. Prior work on this topic has been limited, by (i) providing incomplete guarantees; and (ii) by not addressing GNNs.



Uncovering a previously unknown plateauing-of-robustness phenomenon for different neural network architectures for intermediate levels (between 15% to 30%) of label corruption.

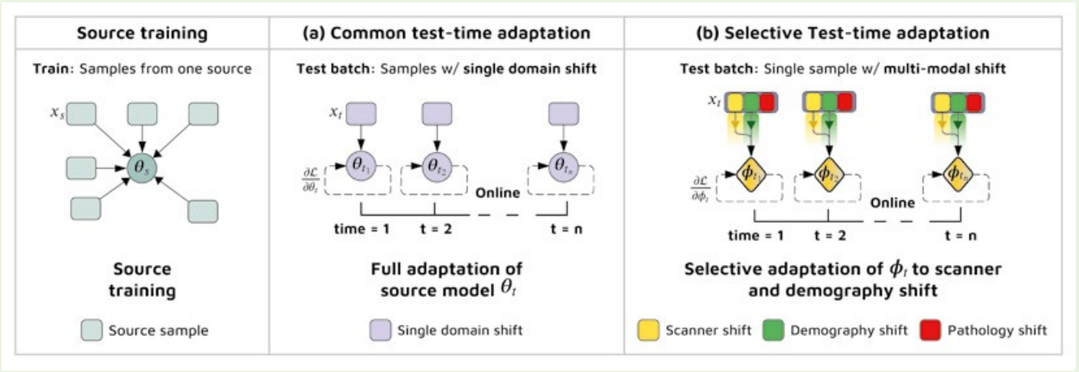
This paper presents the first method for deriving an **exact robustness certificate** against label poisoning for neural networks, including GNNs. The authors approach this by leveraging the **Neural Tangent Kernel (NTK)**, a mathematical tool that describes the behavior of "wide" neural networks (i.e, with a large number of parameters per layer). This allows them to reformulate the complex problem of label corruptions versus learning into a single, solvable optimization problem known as a **Mixed-Integer Linear Program (MILP)**. By solving this MILP, the authors can determine whether a model's prediction for a given data point—or even for an entire test set—is immune to being flipped by label poisoning. The research not only provides a concrete defense mechanism but also offers new insights into how different architectural choices in GNNs affect their robustness, revealing a surprising phenomenon where robustness plateaus at certain levels of data corruption.

This paper's core contribution is the introduction of a method to shift the concept of AI robustness from a "best-effort" practice to a provable property. By providing the first exact certification against a form of data poisoning for neural networks, it provides a clear path toward creating verifiably trustworthy machine learning models that are essential for reliable AI.

Medicine & Healthcare

AI is poised to fundamentally transform medicine and healthcare, offering the potential for earlier, more accurate diagnoses and better-tailored treatments. This revolution will not only lead to improved patient outcomes and increased efficiency but also create significant opportunities for economic growth. This includes applications in **prevention and early detection**, such as AI for wearable devices and screening tools like mammography.

For AI to be successfully deployed in clinical settings, it is essential to develop safe, secure, and trustworthy machine learning techniques. At relAI, our research is dedicated to this challenge, focus-



Selective test-time adaptation. (a) Standard methods fully adapt the trained model to the unseen data. (b) In contrast, this work proposes selective adaptation to unseen data by implicitly excluding pathology shifts.

ing on **robust and data-efficient learning, privacy preservation, and interpretable deep learning.**

The paper "Selective Test-Time Adaptation for Unsupervised Anomaly Detection using Neural Implicit Representations" received a Best Paper Award at the MICCAI Workshop on Advancing Data Solutions in Medical Imaging AI (ADSMI) in 2024.

Ambekar, S., Schnabel, J., Bereca, C. (2024). » Selective Test-Time Adaptation for Unsupervised Anomaly Detection using Neural Implicit Representations. 10.48550/arXiv.2410.03306.

Unsupervised anomaly detection aims to identify unusual data points without a priori knowledge of what constitutes an anomaly. This is particularly challenging in fields like medical imaging or industrial inspection, where anomalies are rare and diverse. The traditional methods often struggle when the distribution of new, unseen data differs from the training data. This paper addresses this by introducing a novel framework for **unsupervised anomaly detection** that proposes a new technique, "Selective Test-time adaptation" with Neural Implicit Representation.

Test-time adaptation during inference allows for adaptation to the unseen data. However, using test-time adaptation on this unseen data for anomaly detection would lead to learning the unseen data distribution, in turn learning the anomalies. This is detrimental, since it skips the process of detecting the anomalies, which involves computing the subtraction between the unseen image and the test-time adapted image. Therefore, this paper proposes to disentangle the distribution shifts that arise due to changes in scanners, demography and pathology. The method then adapts to only the scanner and demography shift while ignoring the pathology shift, which is most helpful for anomaly detection. In contrast to existing works, which focus on adapting the model, this paper proposes to learn the adapted image as a signal with Neural Implicit Representations. Thus, obtaining the selectively adapted image while minimizing the computational cost due to fewer parameters to train.

The paper's core contribution is "Selective test-time adaptation," a robust framework for unsupervised anomaly detection for any trained model. By selectively adapting to new data, it creates a more reliable and trustworthy system that is less susceptible to misinterpreting anomalous data, a critical feature for high-stakes applications.

Robotics & Interacting Systems

The development of autonomous systems, with AI at their core, is unlocking endless possibilities while introducing significant challenges related to **safety, security, and privacy**. Consider the fundamental questions this presents: How can we guarantee the safety of an autonomous agent, like a robot operating in a human environment, when a designer cannot foresee every possible situation? How do we balance the efficiency of AI cloud computing with the heightened risk of security breaches? And how can data be leveraged to adapt to a user's needs without compromising their privacy?

To address these critical questions, relAI focuses on advancing **safe, secure, and privacy-preserving AI** specifically for autonomous agents and interacting systems.

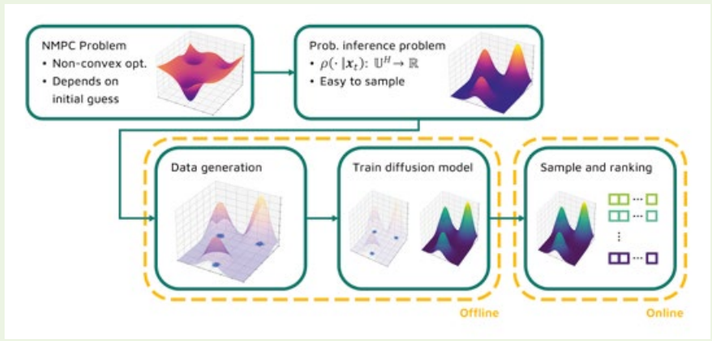
The paper "Toward Near-Globally Optimal Nonlinear Model Predictive Control via Diffusion Models" is the result of research by multiple relAI Doctoral students, supervised by a relAI Fellow. It was nominated for the best paper award and selected for oral presentation at the 7th Annual Learning for Dynamics & Control Conference (L4DC).

Huang, T. Y., Lederer, T., Hoischen, N., Brüdigam, J., Xuehua, X., Sosnowski, S., & Hirche, S. (2025). » Toward Near-Globally Optimal Nonlinear Model Predictive Control via Diffusion Models. Proceedings of the 7th Annual Learning for Dynamics & Control Conference, PMLR 283:777-790, 2025.

Nonlinear Model Predictive Control (NMPC) is a powerful method used in robotics and autonomous systems to make optimal decisions over a future time horizon, all while respecting system constraints. A major challenge with NMPC is its computational complexity; solving the underlying optimization problem can be slow, and the solutions often get stuck in a "local minimum," meaning the system finds a good solution but not necessarily the best one.

This paper presents a novel approach to tackle this challenge by using **diffusion models**, a type of generative AI model typically known for creating high-quality images. Instead of using them for image generation, the authors repurpose diffusion models to learn the distribution of near-optimal control sequences for a given system. By training the model on examples of high-quality control policies, it can then rapidly generate a diverse set of new, feasible control plans. Then, the cost of rolling out these generated plans on the system determines the most optimal control strategy. This method significantly increases the likelihood of finding a **near-globally optimal solution** by exploring a wider range of possibilities compared to traditional single-shot optimization solvers. This approach not only improves the overall performance and efficiency of the controlled system but also provides a more robust and reliable decision-making process.

The main contribution of this work—especially for reliable AI—is an innovative method that improves the performance and consistency of autonomous systems. By using diffusion models to find near-globally optimal control policies, it helps create more predictable and reliable agents that consistently make higher-quality decisions, a vital step for trustworthy AI in critical applications.



Our method proposes a diffusion-based framework to obtain the near-global optima for the NMPC problem.

Algorithmic Decision-Making

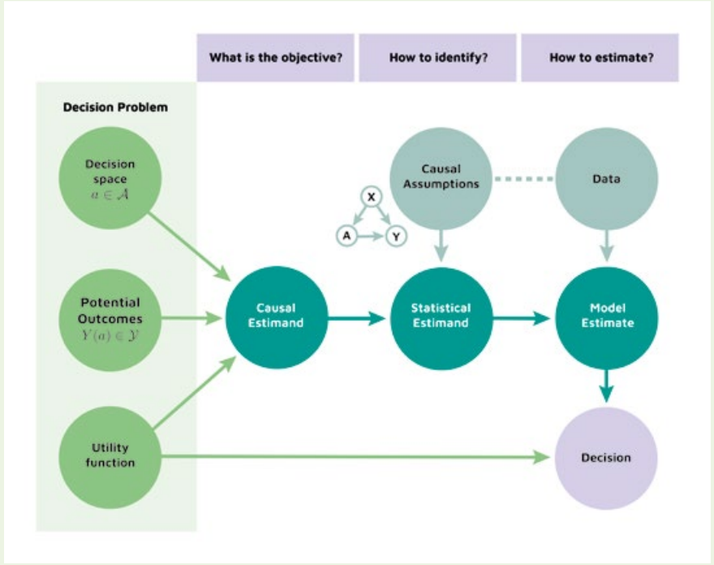
A growing number of AI applications are shifting towards **prescriptive modeling**—systems that not only predict outcomes but also recommend actions in real-world settings. Increasingly, algorithms aid in making decisions that were once reserved for human judgment, such as selecting medical treatments, determining hiring choices or allocating public resources.

At relAI, we develop methodologies for **reliable algorithmic decision-making (ADM)**. This requires addressing core technical challenges, including the absence of a definitive "ground truth" for every prediction and the difficulty of learning from incomplete or selective training data. Importantly, learning to predict well is not the same as learning to decide well, since prediction and decision-making optimize for different goals. Among the open challenges we tackle is the need to reason about counterfactuals—understanding the outcomes that would have occurred under alternative decisions—which calls for methods that incorporate causal reasoning.

Beyond these methodological advances, we study how ADM systems are implemented within **business and public sector institutions**, where reliability depends not only on predictive accuracy but also on how decision systems are embedded into organizational processes. Thinking holistically about downstream outcomes, we examine how algorithms interact with human decision-makers, institutional practices, and policy goals to ensure that automated systems ultimately support trustworthy and effective decision-making.

The paper "Algorithms for reliable decision-making need causal reasoning," a collaboration of several relAI Fellows and Doctoral students, has been published in the prestigious journal Nature Computational Science.

Kern, C., Fischer-Abaigar, U., Schweisthal, J., Frauen, D., Ghani, R., Feuerriegel, S., van der Schaar, M., Kreuter, F. (2025). » Algorithms for reliable decision-making need causal reasoning. Nature Computational Science, 5, 356–360.



This figure illustrates the key components and their relationships in an automated, algorithmic decision-making system. It highlights the flow from defining the problem to using data and causal assumptions to arrive at a solution.

Algorithms for reliable decision-making require **causal reasoning**, especially in contexts such as medical treatment or hiring, where the decisions themselves shape future outcomes. The paper "Algorithms for Reliable Decision-Making Need Causal Reasoning" argues that building ethical and trustworthy decision-making systems demands moving beyond predictive modeling towards a causal perspective. While predictive models can forecast outcomes in static environments, they often fail when their outputs are used to guide actions that, in turn, alter the environment.

The paper shows how a causal framework helps address fundamental challenges in algorithmic decision-making (ADM), most notably the problem of **counterfactuals**, i.e., the hypothetical outcomes of decisions that were not taken. For example, to evaluate a hiring choice, we must consider what would have happened if a different candidate had been selected. By incorporating causal reasoning, algorithms can better handle distributional shifts and estimate the true impact of decisions, rather than relying on correlations alone.

The paper's central contribution is to reframe reliable ADM as a problem of **causal inference** rather than mere prediction. It offers both a compelling argument and a practical roadmap for designing AI systems that are not only accurate but also transparent, fair, and robust—qualities essential for trustworthy deployment in high-stakes applications.

First published in » Nature Computational Science 5, pages 356–360 (2025) by Springer Nature.

Learning & Instruction

BY JOCHEN KUHN AND ENKELEJDA KASNECI

During the past year, relAI has continued to strengthen its education focussed profile. We have introduced "Learning & Instruction" as a new research area and are leading in the field of integrating AI into education.

This research area explores how reliable AI can transform learning and teaching through intelligent tutoring systems, digital learning assistants, and adaptive feedback mechanisms. It addresses the robustness and pedagogical effectiveness of such systems, particularly in the face of incomplete or biased data. Central challenges include ensuring privacy, fairness, and responsibility in data-intensive educational contexts, as well as understanding the broader effects of AI on learners and teachers. A special focus lies on explicability and trust, investigating how transparent and pedagogically meaningful communication of AI-driven actions can foster agency and metacognitive skills. By combining expertise from AI methodology, learning sciences, and education research, this area creates strong synergies with relAI's other domains, ranging from fair assessment and robust learner modeling to socially interactive learning environments and simulation-based training in medicine.

A New Research Area

INTERVIEW: LEARNING & INSTRUCTION – A NEW RESEARCH AREA IN relAI
PROF. DR. ENKELEJDA KASNECI (TUM), PROF. DR. JOCHEN KUHN (LMU)

Q: What will relAI's new Research Area "Learning and Instruction" entail?

The new research area focuses on exploring how reliable AI can be used to transform education in meaningful and responsible ways. It investigates the potential of intelligent tutoring systems, adaptive feedback, and digital learning assistants to personalize learning paths and provide targeted support. At the same time, it examines the broader effects of AI on teaching and learning: how AI systems shape learner motivation, teacher roles, and the dynamics of human-AI collaboration. By bringing together expertise from artificial intelligence, learning sciences, and educational research, "Learning and Instruction" aims to develop robust and trustworthy AI applications that not only advance technology but also serve pedagogical goals and democratic values.

Q: Which aspects of reliability will you address?

Reliability in this context spans several dimensions. First, there is the technical robustness of educational AI systems: ensuring that recommendations, diagnostics, and feedback remain accurate even when data are noisy, incomplete, or biased. Second, there is privacy and security, as educational systems often process highly sensitive information from vulnerable groups such as children or minors. Third, there is the responsibility and fairness dimension: designing learning technologies that protect against discrimination, unequal treatment, or harmful effects on learners. Finally, explicability and trust are central. Both learners and educators must be able to understand and rely on the actions of AI systems. Transparent communication of AI-driven decisions is key to fostering trust, supporting metacognitive skills, and strengthening learner agency. In short, this research area addresses the full spectrum of reliability: safety, security, privacy, fairness, and interpretability.

Q: Why is it important to include education in the relAI curriculum?

Education is both a societal priority and a strategic field of AI application. National and European policy documents emphasize the central role of AI in education, not only to improve participation and equal opportunities, but also to strengthen innovation capacity and democratic resilience. For relAI, this means that research in education is not a peripheral add-on, but a way to directly engage with some of the most pressing societal challenges. By studying how AI interacts with learners, teachers, and institutions, relAI can serve as a "living laboratory" of reliable AI in practice. This allows educational institutions to test and demonstrate AI solutions in its own doctoral programs through adaptive tutoring, personalized feedback, and AI-enhanced learning environments while advancing the field of trustworthy educational technology.

Q: What will be the benefits for relAI students?

The inclusion of "Learning and Instruction" will enrich the academic journey of relAI Fellows on multiple levels. First, students will gain access to cutting-edge research on AI in education, learning directly from interdisciplinary experts at the intersection of computer science, psychology, pedagogy, and ethics. Second, they will benefit from new tools and approaches in their own training, as AI-driven feedback and adaptive tutoring systems are integrated into relAI's graduate education. This means that relAI itself becomes a testbed for innovation, allowing students to experience and critically reflect on AI-enhanced learning firsthand. Third, the new area offers unique opportunities for collaboration across research domains, linking, for example, explainable AI with learner modeling, algorithmic decision-making with fair assessment or robotics with interactive educational systems. Finally, by embedding education into relAI's agenda, Fellows will not only expand their methodological expertise but also develop a stronger awareness of the societal, ethical, and democratic dimensions of AI. This holistic perspective will equip them to become leaders in building the reliable, human-centered AI systems of tomorrow.

Connecting Minds across Borders

Research stays are a crucial component of relAI's strategy to enrich the academic quality and maximize research impact.

VALENTYN MELNYCHUK

From the beginning of February until mid-May 2024, I had the privilege of spending a research stay at the University of Cambridge, where I joined the van der Schaar Lab. This internationally renowned group, led by Professor Mihaela van der Schaar, is one of the world's leading research teams at the intersection of machine learning and healthcare.

During my stay, I collaborated closely with Professor van der Schaar and several of her PhD students, most notably Alicia Curth. Our joint efforts resulted in a major research outcome: the publication of a paper at NeurIPS 2024, entitled "Quantifying Aleatoric Uncertainty of the Treatment Effect: A Novel Orthogonal Learner." As part of our research, we developed a new methodology to quantify the randomness inherent in treatment effects—a challenge that has received little attention in causal machine learning but is of high importance in clinical applications. By proposing a novel orthogonal learner (AU-learner), we were able to provide rigorous theoretical guarantees and demonstrate strong empirical performance across multiple benchmarks.

This collaboration not only strengthened my research profile but also directly aligned with the goals of relAI. One of relAI's central missions is to make AI-based decision-making more reliable, interpretable, and trustworthy. Our paper addresses precisely this: by quantifying aleatoric uncertainty, we provide more robust and transparent insights into causal treatment effects, ensuring that decisions in sensitive domains such as medicine can be made with a clear understanding of potential risks. Thus, the Cambridge stay allowed me to contribute to relAI's overarching agenda while broadening the methodological foundation of my own research.

Beyond publication of concrete research, the stay provided ample opportunities for networking. I participated in numerous events organized by the Cambridge Department of Computer Science and Technology, ranging from research seminars to informal reading groups. These encounters enabled me to connect with leading scientists and fellow PhD students from across Europe and beyond, laying the groundwork for future collaborations. Informal exchanges, such as the lab retreat on the East Coast of the UK, further deepened these connections and provided valuable opportunities for discussion in a relaxed setting.

On a more personal note, this was my first time visiting the United Kingdom, and the experience left a lasting impression. The city of Cambridge itself feels like an open-air museum: every corner is marked by medieval architecture, majestic colleges, and centuries-old libraries. I particularly enjoyed exploring the courtyards and gardens of the colleges, which combine a sense of tradition with a vibrant academic atmosphere.

I also traveled beyond Cambridge, visiting London several times and reconnecting with friends in Glasgow, Edinburgh, and Ireland. Each of these places offered its own charm—London with its cosmopolitan energy, Glasgow and Edinburgh with their blend of history and modernity, and Ireland with its warm hospitality. These trips enriched my stay and gave me a deeper appreciation for the cultural diversity of the region.

Overall, my research stay at Cambridge was an academically productive and personally enriching experience. It resulted in a high-impact publication at NeurIPS 2024, strengthened my ties with leading scholars in the field, and significantly contributed to the goals of relAI. At the same time, it allowed me to discover the cultural richness of the UK and build friendships that extend beyond the academic sphere. I return with not only new scientific insights but also a broader outlook on collaboration, research, and cultural exchange.

CECILIA CASOLO

During my research stay in the fall of 2024, I had the pleasure of being part of the research group led by Caroline Uhler at Broad Institute of MIT and Harvard. My three-month stay in Boston gave me the opportunity to investigate the application of causality and dynamical systems in biology.

The Eric and Wendy Schmidt Center (EWCS) at the Broad Institute is a place where biology and machine learning come together. Here, advances in one field actively drive progress in the other. With the Broad's vast data generation capabilities, EWCS is pushing forward the development of methods that answer fundamental biological questions.

The ultimate goal in biology is to uncover the natural laws that govern life. Achieving this requires a shift from purely predictive models to causal modeling, which can explain mechanisms and make reliable predictions under new conditions. Biology offers a wealth of interventional data—such as gene knockouts and clinical trials—making it an interesting setting for causal inference. Yet major challenges remain: datasets are complex, noisy, and current causal discovery methods have limitations. This is where representation learning becomes powerful. By integrating multi-modal data with domain expertise, it can help identify causal features that remain consistent across different contexts. Such features are essential for reliable downstream tasks, from predicting the effects of unseen interventions to designing better experiments.

A research stay during a PhD is invaluable: it broadens perspectives, builds collaborations, and fosters personal growth. My months in Boston were both scientifically enriching and personally rewarding. I especially recommend visiting Boston in the fall. I had a great time joining the MIT sailing team activities and watching live NBA basketball games. Needless to say, New York City is just a few hours away by bus, and definitely worth a visit! Beyond the lab, the city's vibrant academic and cultural scene made the experience even more inspiring and fun!

Training the **Next Generation** of **AI Leaders**

relAI's educational philosophy is built on the idea that future AI leaders need more than just technical expertise.

Our unique, **interdisciplinary program** is designed to create experts who understand the societal impact of AI and are prepared for diverse careers in both academia and industry.

Hence we provide our students with an entire spectrum of different qualification measures including **scientific knowledge, business expertise, and practical exposure to industry**.

The key components of the educational program are:

Personalized Learning: At relAI, education is tailored to individual needs. Each student crafts a personal development plan (IDP), ensuring their learning journey is uniquely tailored to their goals. A continuous mentorship system, with both academic and industry professionals, provides professional development guidance and sharpens their research focus.

Early Exposure to Real-Life Challenges: relAI prioritizes practical experience. Our curriculum integrates early, hands-on experience through industry internships. This practical exposure bridges the gap between theoretical knowledge and real-world challenges, preparing students for future careers.

International Scope: relAI's program is globally oriented, with a strong emphasis on international exposure. Through a two-way visiting program, our Doctoral researchers gain valuable experience at leading AI centers worldwide. Similarly, many of our Masters students spend time abroad during their studies, with the program supporting their scholarship throughout this period. This global dimension not only broadens students' perspectives, but also enhances their ability to solve AI challenges on a global scale. The diverse international student community further enriches the learning experience.

Focus: the Individual Development Plan

At relAI, we believe in a personalized approach to education, which is why the Individual Development Plan (IDP) is central to our program. The IDP is a tailored concept designed to assist each Masters student and Doctoral researcher in their personal and professional development, aligning with their ambitions and interests.

Students create their IDPs at the beginning of their involvement with the school. The document helps them take stock of their existing skills and experiences, identify gaps they need to fill to reach their goals, and outline their planned activities. This is a collaborative process, as the student works with dedicated mentors (for Masters students) or a Transdisciplinary Thesis Advisory Committee (for PhD students).

The IDP structures a student's journey through three distinct phases:

Orientation Phase: Students evaluate their current skills and identify areas for growth.

Study and Research Phase: The plan specifies academic courses and research activities, including professional development courses and any planned publications or seminars.

Impact Phase: The IDP guides students as they consider future career paths, whether in industry or academia, and identify potential partners for internships or international visits.

The IDP is a flexible document that is revised twice annually to adapt to shifting personal circumstances, academic developments, and professional goals. This ensures that each student receives continuous and personalized guidance throughout their time at relAI.

A Year of relAI Events

Students at relAI benefit from an extensive network and a clear program structure, enhanced by a dynamic calendar of events. These gatherings provide opportunities to learn, meet, socialize, and present their work. Many of these events are organized by relAI staff, but we also actively support and encourage student-led initiatives, fostering a strong, self-driven community.

Scientific Advisory **Board Meeting**

On October 30, 2024, the relAI Scientific Advisory Board (SAB) met in Munich/Garching, where SAB Members spoke with relAI Directors and Fellows, students, and the Management Team to **gain insights into the relAI program from various perspectives within the relAI family**. The SAB advised relAI on the future directions of the program, suggesting strategies for optimizing internal structures and expanding our network.

Retreat

The annual retreat, held for the fourth time in 2025, was **a major highlight of the relAI year**. It fostered intense scientific discussion across various aspects of reliable AI and strengthened the bonds within the community.

Collab Accelerator

The first relAI Collab Accelerator kicked off in November 2024. **This workshop fostered stronger collaboration and research connections** within the relAI family, especially among relAI students.

Student **Socials**

Our students regularly meet for **various social events**, such as at game nights, our Christmas party or on hiking tours together.

Welcome Days

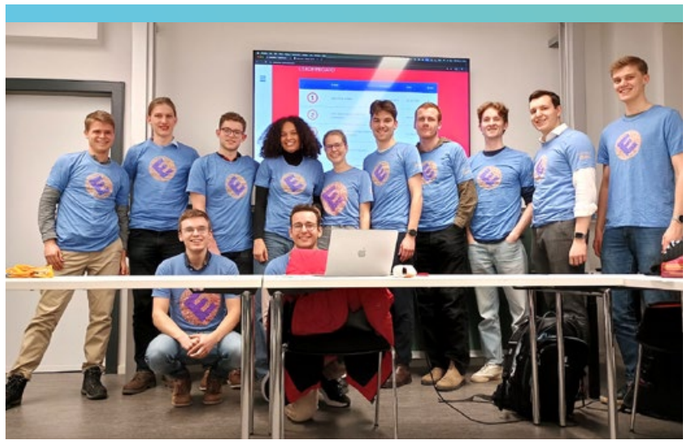
At the start of the fall semester, our Welcome Day provides **a warm introduction for the new cohort**. This event connects new and returning members in an energizing social gathering, giving them a chance to get acquainted with the relAI community and its mission.

TUM Open Day

In October 2024, the Konrad Zuse School was represented at the TUM Open House on the Research Campus in Garching. At our booth, we demonstrated how AI works through engaging games and simulations, which was a **great opportunity to share our vision of reliable AI and to connect with students, researchers and the public**.

Estimathon

In January 2025, a team of 13 talented Masters and PhD **students from relAI showcased their quantitative skills and teamwork** in an exciting Estimation Competition. The participants had 30 minutes to work on 13 estimation challenges, such as "What is the average discharge of the Isar when it meets the Donau in m^3/s?"



AI-HUB@LMU Grand Opening

relAI supported the organization and joined the opening ceremony of our partner AI-HUB@LMU to celebrate its founding. **AI-HUB@LMU is a platform that unites all 18 faculties of Ludwig-Maximilians-Universität München** into a joint scientific community and aims to advance research, teaching, and transfer in artificial intelligence and data science at LMU.

International Summer School

End of July 2025, relAI was delighted to **welcome a group of Chinese students** for the first relAI International Summer School. With the engagement of relAI Fellows and PhD students, the school promoted an educational exchange focused on the development of reliable AI, with the goal of exploring current trends in AI development.

Career Fair

In October 2025, relAI, MDSI, MCML, and the AI Hub@LMU hosted the **first Munich Career Fair for AI and Data Science**. The event brought together 11 industry partners and over 150 students, offering networking opportunities and presentations from the partners about their work and job opportunities in the fields of AI and data science.

Munich AI Lecture Series

The Munich AI Lectures are a collaborative effort by all major AI initiatives in Munich, coordinated by baiosphere. Held approximately **once a month**, **these lectures feature insights and ideas from experts**, fostering a shared platform for knowledge exchange.

DataFest 2025

In March 2025, relAI supported the 8th edition of DataFest Germany. The event is **an annual data-driven competition**, **commonly referred to as a "hackathon,"** that alternates between Mannheim and Munich. It is organized in collaboration with partners from industry and research institutions. A team of relAI students participated in this competition.

Student Seminar

In a **monthly** seminar organized by the student representatives, our students not only present their work, but also engage in discussions and a social get-together.

Each year, the three Konrad Zuse Schools **come together** to enhance their collaboration and discuss emerging challenges. relAI proudly hosted this year's Zuse School Meeting **at the TUM campus in Garching** in October 2024.

Zuse School Meeting

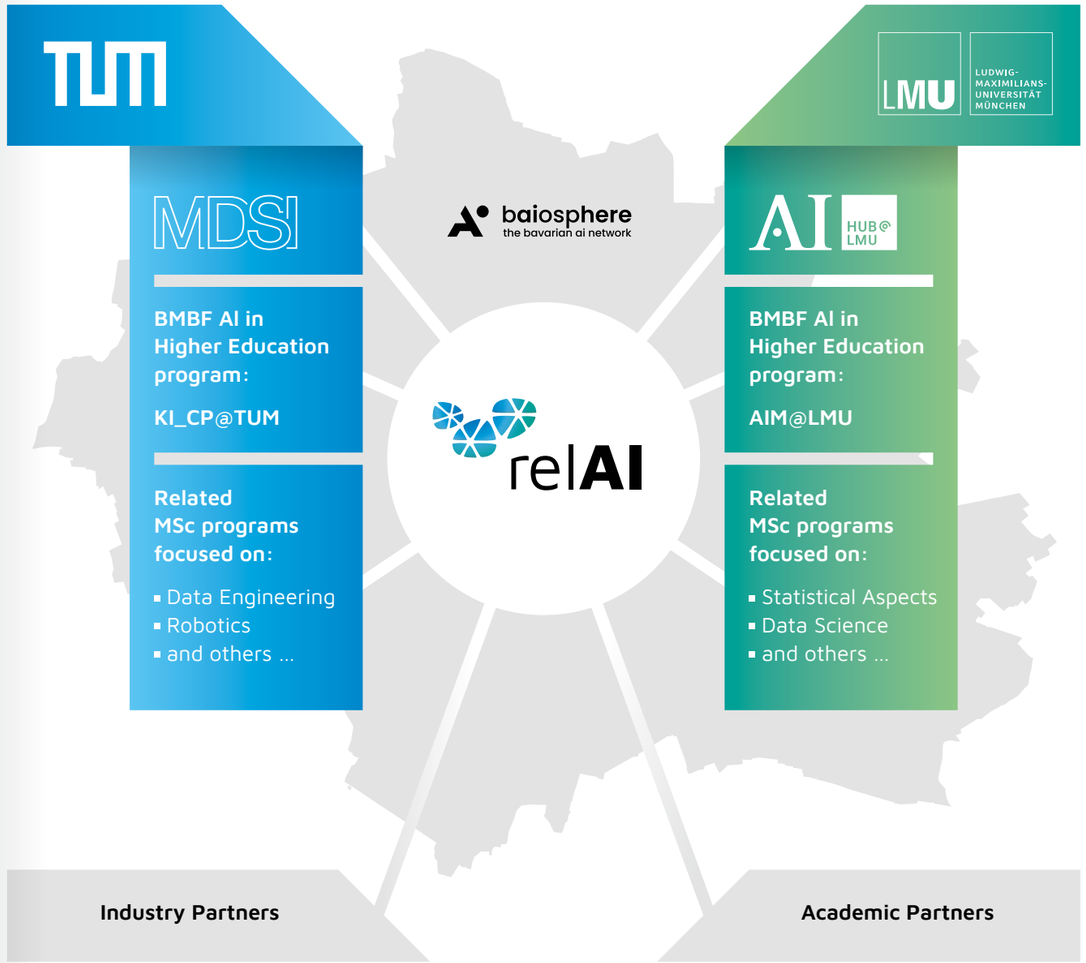
Building Strong Connections

Our relAI program is built on a robust network that spans international academic partners and leading industry experts. This strong foundation creates a dynamic ecosystem where researchers, students, and partners can thrive. In addition to these internal connections, relAI is embedded in the baiosphere, the vibrant, multifaceted AI ecosystem of Bavaria and particularly Munich, which integrates research, education, and applications across various domains.

We are a key link within this local network, bringing together existing institutions through our Research Areas. For example, we connect with the Munich Data Science Institute (MDSI) and the Munich Center for Machine Learning (MCML) in the field of Mathematical and Algorithmic Foundations. In Medicine and Healthcare, we work with LMU Klinikum and the Klinikum Rechts der Isar. In Robotics and Interacting Systems, we collaborate with the Munich Institute for Robotics and Machine Intelligence (MIRMI), while our work in Algorithmic Decision-Making links us to the Munich Center for Mathematical Philosophy (MCMP).

This collaboration extends to education as well, where we build bridges to programs such as KI_CP@TUM and AIM@LMU, as well as various MSc programs in Data Engineering, Robotics, and Data Science.

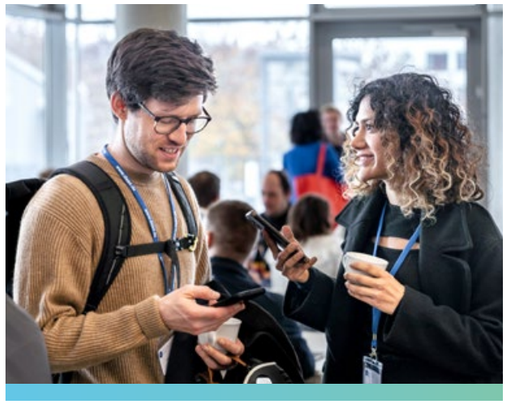
»Our research projects are **bringing together** professors, post-docs, Doctoral candidates, and MSc students across various academic departments and research centers.«



Engaging with the Public

Public communication plays a vital role at our school. We share the latest advancements in reliable AI and keep our audience informed about our activities through our website and social media channels.

WWW X LINKEDIN BLOG BLUESKY



Our **website provides a comprehensive overview** of the school's research, education, and governance structure. It reaches a diverse audience, including potential applicants, relAI members, and the general public, attracting an average of over 3000 unique visitors each month.

Social Media channels and the news section on our website serve as our primary communication channels to rapidly disseminate information. relAI actively maintains three social media channels: LinkedIn and X (formerly Twitter) as well as Bluesky, with more than 185000 impressions within the last year. DAAD supports our outreach activities by sharing relAI updates through its online channels.

Students at relAI are actively involved in the relAI Blog, a platform for sharing innovative research and developments from our school. The blog features posts on introductory research topics, the latest project outcomes, and insights into relAI life.

Furthermore, relAI provides opportunities for personal and direct engagement by participating in **outreach events**, such as Open House events at the universities.

Working towards a successful future

The long-term importance of AI, and its reliability in particular, is undeniable. Our goal is to establish the Konrad Zuse School as a lasting institution in this critical field.

To ensure the long-term sustainability of the school, we have implemented quite a number of initiatives:

Fostering the relAI family
We actively work to foster a strong sense of community among our members, creating what we call the "relAI Family." As highlighted in this report, a wide range of events regularly brings everyone together, from students to Fellows and our valued partners.

Cultivating an Alumni Network
With the support of Fellows and the management team, a group of students who are about to graduate has begun to establish a framework for our alumni network. This initiative aims to integrate former members into the relAI family, for instance by sharing first-hand accounts of their transition from academia to industry.

Maintaining High-Quality Publications
As detailed in the research section, relAI members have published a significant number of high-quality papers. This demonstrates that our research output is well-received within the academic community, and our members will continue to strive for this high standard.

Expanding Our Network
relAI receives numerous applications from academic researchers and industry/academic partners alike who are interested in joining our network in order to contribute to as well as benefit from our research and education program.

Engaging in Science Communication
A new science communication course has been designed for the relAI education program, providing our students with key skills in communicating complex research to a broader audience. While students regularly contribute to the relAI blog by writing about their research, we actively promote this content—and other project news—through social media channels to ensure our work reaches a wider public.

Expanding Application Areas
As AI technology is adopted in an increasing number of fields of application, relAI will continue to evaluate and add new application areas. As outlined in this report, relAI Fellows Prof. Kuhn and Prof. Kasneci are currently in the process of establishing Learning & Instruction as a new Application Area.

The numbers behind relAI

2022

relAI was founded in 2022 and welcomed its first students in the same year. It celebrated its grand opening a year later in 2023.

relAI students have been recognized on **12 occasions at prestigious conferences** such as NeurIPS, ICLR, and ICML. Their achievements include an IJAR Young Researcher Award, an Outstanding Paper Award, three Best Paper Awards, and one Runner-Up Award. Additionally, six

12

articles were selected for oral and spotlight presentations, ranking among the top 3%.

The social media channels of the graduate school generate **more than 2700 followers** across the platforms LinkedIn, X (previously Twitter) and BlueSky. Following relAI on social media is a great way to stay updated about the school! The accounts feature event invitations, awards, research updates and other news about the school.

2.700+

3000+

The relAI Blog, a platform to share cutting-edge research with academics, industry professionals and the general public, started in 2024. Its 11 postings to date receive an average of **over 3000 visitors per month.**

32

The international character of the school is reflected in the fact that **our relAI students represent 32 different countries.**

40 Mio

The Zuse schools were established as part of the AI action plan of the German Federal Ministry of Education and Research (BMBF-Aktionsplan Künstliche Intelligenz). **Together, all Zuse Schools receive €40 million in funding.**

The members of the Zuse school have published their research in **over 180 publications**, including prestigious venues such as NeurIPS, CVPR, and Nature.

180+

45

relAI Fellows are holders of **45 prestigious distinctions and grants**, including 22 European Research Council (ERC) grants, 5 Gottfried-Wilhelm-Leibniz awards, and 2 Alexander von Humboldt Professorships.

There are three Konrad Zuse School of Excellence, all funded by DAAD. Our "sister" Zuse Schools are ELIZA and SECAI. ELIZA is closely connected to the European Laboratory for Learning and Intelligent Systems (ELLIS), a leading European network for AI research. SECAI stands for School of Embedded Composite Artificial Intelligence and is based at the Technical University Dresden and University of Leipzig.

3

92

In summer 2025 a total of **92 MSc and Doctoral students** formed the heart of relAI. The school has attracted top international talent and through its structure and community, it supports exceptional young minds in their education, combining technical brilliance with awareness of the importance of AI's reliability.

48 fellows conduct courses for the students, provide supervision of the doctoral researchers, and jointly conduct research together.

48

13 Mio

The budget allocated to relAI is €13 million, ramping up from €0.6 million in 2022 to €3 million in 2025.

22 companies partner with relAI. Close ties to industry emphasize the transdisciplinary character of the school. The industry partners engage with the students in networking events, discussing application-focused challenges. Furthermore, representatives from the industry partners and other associated companies serve as mentors on the transdisciplinary thesis advising committee (TTAC) of the relAI Doctoral students.

22

Among all students funded by relAI, **a total of 35% are female**, highlighting relatively high gender diversity as compared to the trend in the field.

35%

1300+

As of July 2025, **more than 1300 applicants have applied to relAI**. Of these, around 85% come from countries outside of Germany, which showcases the international focus of the school.

20+

More than 20 events took place in the academic year 2024–2025. relAI events include talks given by world-class experts, an international summer school, community events such as the Welcome Days, our retreat and even student-organized seminars and a collaboration workshop.

40+

relAI students have gained **valuable experience** through more than 40 research visits, industry internships, project groups, and guided research of the relAI curriculum. These opportunities offer early exposure to real-world challenges and facilitate the transfer of research findings.

Beyond the Lab: Our Commitment in Action

Our **commitment to reliable AI** extends beyond foundational research.

We bring theory to life by exploring **real-world applications**, expanding **application areas**, collaborating with **industry leaders**, and tackling **everyday challenges**.

AI, Reliability, and the Future of Insurance

AN INTERVIEW WITH MUNICH RE EXPERTS: CHRISTIAN HOBELSBERGER, INTEGRATED STUDENT, RELAI MSC STUDENT, MARTIN THORMÄHLEN, IT PROGRAM MANAGER & LEAD OF THE TECH TREND RADAR, AND ANDREAS NAWROTH, LEADING EXPERT AI & QUANTUM



Artificial intelligence is reshaping the insurance industry by driving new capabilities in automation, decision-making, and risk analysis. Munich Re, a global leader in (re)insurance, is at the forefront of this evolution and a committed partner in relAI's mission to develop responsible and reliable AI. With its deep roots in data-driven risk assessment and longstanding focus on emerging technologies, Munich Re offers a valuable industry perspective on the future of AI.

We've asked Christian Hobelsberger, a relAI M.Sc. student and integrated M.Sc. student at Munich Re, together with Martin Thormählen, IT Program Manager & Lead of the Tech Trend Radar, and Andreas Nawroth, Leading Expert AI & Quantum, to share their key insights from the Tech Trend Radar 2025 and reflect on how collaborative research and technological foresight are shaping the future of insurance.

Q: How does Munich Re evaluate the relevance and potential impact of technological advancements on the (re)insurance industry as they accelerate?

Martin Thormählen: At Munich Re, we've systematically monitored and assessed tech trends for over a decade through our annual Tech Trend Radar report, which we developed in close collaboration with our subsidiary ERGO. This strategic tool helps us identify developments that could have a lasting impact on internal operations and the broader insurance ecosystem.

We examine trends from a business-first perspective. For instance, the 2025 Radar identifies Generative AI, AI agents, and spatial intelligence as particularly relevant in the current cycle. These technologies touch on everything from risk modeling, underwriting to claims. They're no longer speculative; but are beginning to transform how we assess and manage risk. This kind of technological foresight is essential in reinsurance because understanding emerging systemic risks is part of our core mission.

Q: What's the current status of AI, and how could it affect the insurance sector?

Andreas Nawroth: AI has made remarkable progress in recent years. We've progressed from narrow AI to large language models (LLMs), which can reason and solve complex problems. In fact, two AI systems recently won gold medals in mathematics competitions, which is a clear sign of how far the technology has come.

But why does this matter for insurance? Our industry relies heavily on actuarial mathematics for pricing and risk assessment. We're exploring how good these models are in general and how well they understand insurance-specific knowledge. We're also investigating where they can be applied effectively.

At the same time, reliability and explainability are essential in insurance. Decisions must be transparent and trustworthy, so we must ensure that AI systems meet these standards before integrating them into critical processes.

Q: Christian, you're involved with both relAI and Munich Re. How does that influence your work and research?

Christian Hobelsberger: It's a unique opportunity. I'm currently working on a joint research project that focuses on evaluating and reducing uncertainty in large language model (LLM) outputs. This is crucial for building trust in AI-generated information. This is particularly relevant in insurance, where decisions often have significant financial or legal implications.

The experience is valuable because of the dual perspective. relAI provides me with deep scientific tools and methods, and working at Munich Re challenges me with real-world complexity and application needs. This interaction ensures that the research remains both grounded and impactful.

Q: AI Agents are a prominent trend in the Tech Trend Radar 2025. How is Munich Re approaching their use?

Andreas Nawroth: We see AI agents as a step up from mere interaction to actual action, which makes them highly relevant to businesses. These agents can automate multi-step processes such as product design and market analysis, creating significant economic value.

However, reliability becomes critical when agents start taking actions. In insurance, for example, decisions must be transparent and trustworthy. That's why we're carefully assessing the maturity of these capabilities for different processes, identifying areas where AI agents can add value, and determining how to ensure the level of reliability required by our industry.

Munich Re is among the first companies to explore this from both an operational and a business opportunity perspective. With a dedicated team and the aiSure product, we're developing solutions to insure AI systems themselves since 2018, addressing the new risks that come with these technologies.

Q: Although Spatial Intelligence isn't always in the spotlight, you've identified it as a critical trend. Why is it important now?

Martin Thormählen: The world is becoming more connected—and more volatile. Many of the risks we address, such as wildfires and floods, are inherently spatial. Thanks to advances in geospatial artificial intelligence (AI), satellite data, and drone imagery, we can now assess risks and damages much more dynamically and precisely. In the future, large world models, which aim to simulate real-world physics and dynamics, could become even more important. This will have a significant impact on underwriting and claims. After a catastrophe, for instance, spatial data enables us to triage claims more quickly, target response efforts, and improve pricing accuracy.

However, the relevance extends beyond natural catastrophes. Spatial intelligence which leverages emerging World Models have the potential to optimize safety of autonomous vehicles and therefore significantly influence motor insurance pricing. Also humanoid robots will be able to tremendously improve with World Models, that allow them to understand and interact with the real world. This trend may not yet receive as much media attention as AGI or generative AI, but it has immense practical value for insurers.[AS1]

Q: Christian, how has your view of what it means for AI to be reliable changed because of relAI? How does Munich Re benefit from this collaboration?

Christian Hobelsberger: relAI has taught me that reliability extends well beyond accuracy. It's about transparency, robustness under uncertainty, and the long-term behavior of systems. This mindset is essential in insurance, where even minor mistakes can have significant consequences.

At Munich Re, I see this philosophy being put into practice every day. Models aren't just evaluated based on benchmarks; they're also tested on how well they integrate into human workflows, how understandable their decisions are, and how they perform in edge cases. The alignment between academic training and real-world application is one of the partnership's most valuable aspects.

In my view, relAI provides scientific depth and the best possible learning environment, while Munich Re adds context, responsibility, and impact. The knowledge I've acquired is not only applied in practice, but also valued by my colleagues. This combination makes our collaboration unique.

For more on these trends, see the full Tech Trend Radar 2025, developed by Munich Re and ERGO IT Strategy, available on Munich Re's website.

Student Experiences

Internship with BMW Digital Campus Munich

OMAR BOUATTOUR,
MASTER STUDENT
FROM TUNISIA

Omar Bouattour, a Software Engineering student from Tunisia, is about to finish his Elite Masters and graduate from the relAI program. As part of his academic curriculum, he completed a 5-month internship at BMW Group, within the Digital Campus Munich (DCM). Here is what he tells us about his experience:

Assigned to the Data Engineering Team, I had the opportunity to work on an independent project while collaborating closely with experienced professionals and fellow interns.

Onboarding and Work Environment

My first day began with a warm welcome from my mentor, who met me at the entrance, introduced me to the campus, and guided me through my hardware setup. I was also enrolled in BMW's orientation and training programs, which provided an overview of the company's digital initiatives and workflows.

The Data Engineering Team consisted of two full-time engineers and two interns, fostering a collaborative yet focused environment. My schedule was flexible: although I typically worked seven hours per day, I had the freedom to manage my time as long as my tasks were completed. Daily one-on-one meetings with my mentor ensured continuous guidance and feedback. Lunch breaks were often shared with the team at the BMW Mensa, which helped strengthen personal connections. Social highlights included visiting the Weihnachtsmarkt together and attending BMW's Christmas party, offering memorable cultural and networking experiences.

Project and Responsibilities

My primary assignment was to develop a machine learning model capable of detecting anomalies in vehicles using multi-series historical data transmitted by connected cars during their online sessions. This project required designing, implementing, and evaluating multiple approaches to determine the most effective solution.

Technically, I worked extensively with AWS services, including Athena for querying large datasets and SageMaker for training and deploying models. I experimented with Hugging Face transformer models as well as LSTM-based auto-encoders for anomaly detection. The task involved handling vast, messy datasets, performing preprocessing, and tuning models for optimal performance. At the conclusion of the project, I presented a comparative analysis of the implemented approaches, highlighting their respective advantages and limitations.

Reliability, Safety, and Trust

My work on anomaly detection was directly tied to the core of relAI's mission: building reliable AI systems. In the automotive industry, accurately and promptly identifying anomalies in vehicle data is critical for safety and security. By developing a model that can find subtle deviations from normal operational behavior, I contributed to BMW's efforts to enhance fleet safety. The model provides a crucial layer of trust by continuously monitoring vehicle health and flagging potential issues before they can escalate.

Skills and Learning Outcomes

This internship allowed me to bridge the gap between academic machine learning knowledge and its application in enterprise-specific contexts. I gained hands-on experience in:

- Managing and processing large-scale, imperfect datasets.
- Designing and evaluating deep learning architectures for time-series anomaly detection.
- Utilizing cloud-based tools to deploy data-driven solutions.

Beyond technical skills, I learned how to navigate corporate culture, understand organizational hierarchies, and communicate effectively with both colleagues and upper management. I developed valuable networking skills and built professional relationships that I expect to maintain in the future.

Memorable Moments and Conclusion

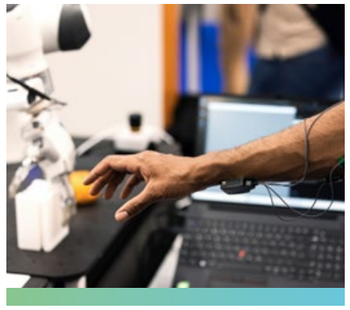
One of the most rewarding moments was the conclusion of my internship, when I was presented with a certificate of completion accompanied by positive remarks on my performance. My team also surprised me with a thoughtful farewell gift: a Mini Cooper coffee mug and a BMW 5 Series miniature model, both symbolic of my time at BMW.

Overall, my internship at BMW was a formative experience, combining technical growth, cultural exchange, and lasting professional connections. It deepened my expertise in machine learning applications for connected vehicles and strengthened my confidence in pursuing a career at the intersection of AI and automotive innovation.

Science communication activity at 2025 automatica fair

Ahmed Abdelrahman is a relAI PhD student from Sudan, working on neuromorphic methods for reliable myoelectric control of robots and neuroprostheses at the Munich Institute of Robotics and Machine Intelligence (MIRMI). He shares his experience in science communication as an exhibitor at the biennial automatica 2025 robotics trade fair:

AHMED ABDELRAHMAN, PHD STUDENT FROM SUDAN



At the automatica 2025 fair, I led a demonstration of myoelectric robot control for object manipulation, where a robot is partially controlled through human intention by decoding their electrical muscle activity. To perform tasks like moving objects and folding cloths on a table, a user controls the grasping actions of a robot manipulator using electromyography (EMG) signals from non-invasive sensors placed on their forearm. By recognizing distinct patterns in the user's muscle activity that encode their intention to grasp or release objects, a machine learning model transforms these intentions into commands for the robot to execute. The project is a part of my relAI PhD research into reliable, AI-based robot control from biological signals, such as for neuroprosthetic devices that can restore lost limb function.

This event was a great opportunity to showcase my work to the general public and discuss its implications with visitors of varying backgrounds, experiences, and opinions. Due to the diversity of the audience, I learned to quickly adapt my presentations according to the attendees' level of expertise. Overall, the challenge of succinctly explaining my research to visitors with no prior knowledge and to experts alike has greatly aided me in rethinking and improving my approach to effective scientific communication.

Get involved

(Prospective) students:
We usually publish our call for MSc applications in the spring and our call for Doctoral students with excellent transcripts in late fall. You can join our mailing list to get informed about the start of the application period. We expect a Bachelors (and MSc) degree (or equivalent) in computer science, mathematics, engineering, natural sciences, or other data science/machine learning/AI related disciplines, a high grade point average, and a genuine interest to work on a topic of reliable AI.

Industry partners:
We are always looking for industry partners working on reliable AI. Please reach out at coordinators@zuseschoolrelai.de to discuss a collaboration.

Politics, press and general public:
You can contact us via our general mail: coordinators@zuseschoolrelai.de.

LEGAL NOTICE

Responsible party in terms of press law:

relAI Directors

Prof. Stephan Günnemann
Munich Data Science Institute (MDSI)
Technische Universität München
Walther-von-Dyck-Straße 10 · 85748 Garching · Germany

Prof. Dr. Gitta Kutyniok
Ludwig-Maximilians-University München
Akademiestraße 7 · 80799 Munich · Germany

Editorial Team
Dr. Mónica Campillos, Daria Kozlova, Nora Ott, Dr. Andrea Schafferhans

Graphic design
grafikcafé :: feines design

Photo credits:
The photos were provided with the kind permission of:
TUM/Astrid Eckert, TUM/Andreas Heddergott, Conny Mirbach, Ansgar Pudenz, DAAD / Siegfried Michael Wagner, MIRMI/TUM, Alessandra Wolfsberger (Unsplash)

relAI
Konrad Zuse School
of Excellence in Reliable AI

Munich Data Science Institute (MDSI)
Walther-von-Dyck-Str. 10
85748 Garching, Germany

info@zuseschoolrelai.de
www.zuseschoolrelai.de

With funding from the:

Federal Ministry
of Research, Technology
and Space

Zuse Schools
Konrad Zuse Schools of Excellence
in Artificial Intelligence