Press Release

Partnership combines AI testing with robust evidence

October 12, 2026

TÜV SÜD and QuantPi are joining forces to ensure AI safety

Munich/Frankfurt am Main. TÜV SÜD and QuantPi are to collaborate on testing and assessing AI systems. As TÜV SÜD's first CARAT partner laboratory worldwide, QuantPi will contribute its technology for statistically validated AI testing to the partnership. Together, the companies are thereby creating a structured pathway to verifiable test evidence for the conformity assessment and certification of AI-based products – particularly in the fields of Physical AI and autonomous machines.

Regulatory requirements meet a new technological reality

The EU AI Act and the new Machinery Regulation (EU) 2023/1230 mean that the requirements for demonstrating the safety of AI-based functions are increasing. Manufacturers must demonstrate that AI systems function reliably, particularly where AI is used as a safety component or performs safety functions. The EU AI Act sets out corresponding requirements for conformity assessment of AI systems in regulated products.

A key challenge here lies in generating robust and traceable test evidence. Currently, there are only a limited number of established procedures for AI-based systems that make test results usable for conformity assessment with statistical validity, reproducible methodology, and end-to-end traceability. At the same time, international standards for integrating artificial intelligence and functional safety are being developed further. These include the ISO/IEC 22440 series, which sets out requirements and guidelines for AI in safety-critical applications.

Statistical evidence instead of mere benchmark values

This is where the collaboration between QuantPi and TÜV SÜD comes in. The QuantPi platform can describe the behaviour of AI systems with statistical guarantees and defined confidence limits. This provides quantifiable and reproducible information about the performance of an AI system that goes beyond traditional benchmark results or qualitative assessments.

The method is model-independent and does not require access to internal model parameters or training data. QuantPi operates the platform in accordance with an ISO/IEC 17025:2017-aligned quality management system. This ensures that the processes and results produced can be traced and audited. Thus, the combination of statistical validity and an auditable testing process creates a robust basis for using the test results as evidence in the context of conformity assessment and certification.

“QuantPi has demonstrated its ability to deliver high-quality and scientifically robust results for AI testing. The company therefore complements our AI testing and certification programme perfectly,” says Benedikt Pulver, Head of Machine Safety at TÜV SÜD. “Through this collaboration, we are combining our respective areas of expertise to assess even highly complex AI systems and to provide our customers with a clear path towards future certification and conformity assessment.”

Philipp Adamidis, CEO and co-founder of QuantPi, adds: “Companies developing AI-based products and services for regulated markets and safety-critical applications often face the same challenge: their systems function as intended, but there is a lack of robust evidence that meets the requirements of a certification body. With the formal recognition of QuantPi’s evidence by a Notified Body responsible for machinery regulation, companies now have a clear and practical way to systematically test AI systems and utilise the results for conformity assessment.”

Physical AI poses particular challenges for AI safety

The new testing methodology is particularly relevant for Physical AI, autonomous machines and robotics. Humanoid robots, autonomous mobile robots (AMRs) and other AI-based machines operate in dynamic environments and interact with people. It is difficult to model and test their potential operating environment in full.

Statistical methods can provide a valuable foundation for drawing reliable conclusions about a system’s behaviour from a defined number of tests. This is becoming increasingly important as autonomous systems become more widespread.

At the same time, new international approaches are emerging to systematically integrate artificial intelligence and functional safety. The ISO/IEC 22440 series, which is currently being developed, covers requirements and guidelines for using AI in safety-critical functions, among other things.

“The shift towards Physical AI is fundamentally changing safety requirements. AI systems are increasingly taking on tasks that are directly linked to the physical world,” says Benedikt Pulver. “It will therefore be crucial for manufacturers to be able to demonstrate safety in a measurable, traceable and robust manner throughout the entire assessment process.”

Joint framework for AI testing and certification

Through their collaboration, QuantPi and TÜV SÜD are combining their respective expertise to create a comprehensive framework for the efficient, robust, transparent and traceable assessment of AI systems. The framework supports both the structured conduct of AI tests and the preparation of results for subsequent conformity assessment and certification.

The initial focus is on Physical AI and regulated industrial applications. In this area, the market’s increasing readiness for AI-based products and robotics is met with particularly high safety and evidence management requirements.

This collaboration is also embedded within the growing international Physical AI safety ecosystem. Both TÜV SÜD and QuantPi are members of the NVIDIA Halos ecosystem. Halos is a comprehensive safety platform for physical AI, and through the AI Systems Inspection Lab it provides an ANAB-accredited inspection programme for AI functional safety in physical AI. The aim is to standardise the assessment of safety systems and pave the way for independent body testing.

This creates an increasingly integrated ecosystem: AI systems are developed and systematically tested, and the results are statistically validated and prepared for subsequent independent assessment.

Trust in AI is becoming measurable

As AI becomes more widely used, the issue of safety and trust is shifting towards the technical infrastructure of products. Therefore, it is becoming increasingly important for manufacturers to be able to demonstrate the performance and behaviour of their AI systems using robust data. TÜV SÜD and QuantPi are developing a concrete approach to this. AI safety can be measured and traced, making it more applicable in regulated markets. This could be particularly significant for European manufacturers looking to bring AI-based products to international markets safely, compliantly and scalably.

About QuantPi

QuantPi is the validation layer for enterprise AI. The QuantPi platform proves what an AI system actually does - any model, any modality, agentic or physical - without ever seeing the model, and quantifies how much to trust the answer. No leaked benchmarks, no saturated leaderboards: statistical confidence intervals on every result, generated inside the customer's own environment. Spun out of the CISPA Helmholtz Center for Information Security and in partnership with the world's leading TIC organizations. quantpi.com/

Further information:

- tuvsud.com/en/topics/artificial-intelligence
- tuvsud.com/en/services/product-certification/ai-compliance-mark-for-low-risk-ai-systems

Note for editorial staff: The press release is available at: tuvsud.com/newsroom.

Media Relations

TÜV SÜD Public Relations Westendstraße 199 80686 Munich, Germany	Dirk Moser-Delarami Phone +49 89 5791-1592 Email dirk.moser-delarami@tuvsud.com Internet tuvsud.com/newsroom

Founded in 1866 as a steam boiler inspection association, the TÜV SÜD Group has evolved into a global enterprise. More than 30,000 employees work at over 1,000 locations in about 50 countries to continually improve technology, systems and expertise. They contribute significantly to making technical innovations such as Industry 4.0, autonomous driving and renewable energy safe and reliable. tuvsud.com