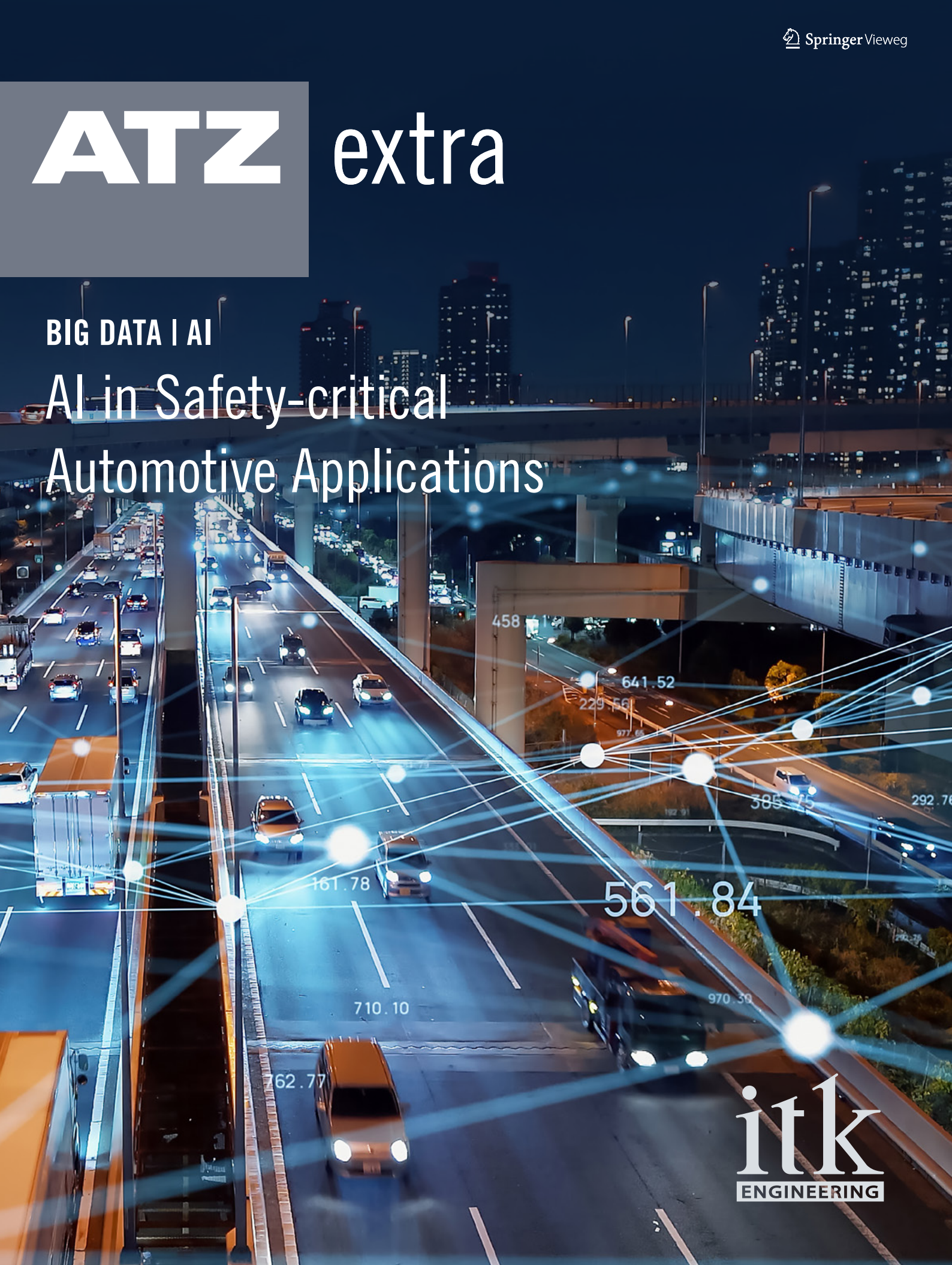
Springer Vieweg

ATZ extra

BIG DATA I AI

AI in Safety-critical Automotive Applications



itk ENGINEERING

AI in Safety-critical Automotive Applications

Artificial Intelligence (AI) is taking the development of vehicles and their functions to a new level. ITK Engineering is dedicated to the question of which analysis methods and systematic approaches can be used to ensure the necessary safety aspects of machine learning systems to comply with ISO PAS 8800.

Rapid advances in artificial intelligence are affording engineers myriad opportunities to develop innovative automotive features. The downside of these advances is that they put automakers under tremendous pressure to innovate. Accidents involving automated vehicles are another issue [1]. What's more, governments are rushing to regulate AI [2]. All this gives rise to a key question: What could a standardized approach to developing safety-relevant AI systems in vehicles look like? The ISO PAS 8800 specification is going to furnish this framework for developing these systems in vehicles. This standard will guide manufacturers and suppliers that develop AI systems for automotive functions [3].

THE CHALLENGES OF DELIVERING SAFE AI SYSTEMS

AI systems typically serve to solve challenging problems. In vehicles, they primarily enable automated driving functions. Deep learning systems are crucial to this end. They use algorithms to analyze vast quantities of data to learn complex tasks, for instance, how to recognize objects. This transpires in the open world, so engineers have to find a way to continuously develop and assess the safety of AI systems in the vehicle. Another challenge to consider is the very nature of AI-driven methods. For one, deep neural networks are nonlinear and exceedingly complex, which makes them difficult to comprehend and explain. For the other, their limited generalization ability and robustness narrows their functional range. Data-centered tasks such as collection, annotation, and dataset design merit special consideration because the behavior of AI models based on Machine Learning (ML) is implicitly specified and determined by the data used.

WRITTEN BY

Dr. Stefan Held
is Lead Engineer at ITK Engineering in Holzkirchen (Germany).

Andreas Bossert
Is Senior Expert Engineer Verification and Validation at ITK Engineering in Holzkirchen (Germany).

Dr. Frank Lenzen
is Expert Engineer Functional Safety at ITK Engineering in Rülzheim (Germany).

Dr. Ulrich Sutter
is Senior Manager Functional Safety at ITK Engineering in Rülzheim (Germany).

SAFETY ARGUMENTATION FOR AI SYSTEMS

Constructing an argumentation, referred to as ‚assurance argument' to substantiate the product's safety is integral to every safety standards-compliant development project. In this context, "safety" means that the chances of personal injury attributable to a system malfunction or to reasonably foreseeable misuse has been reduced to an acceptable level of risk. The assurance argument has to be structured systematically [4]. Engineers can render it schematically using Goal Structuring Notation (GSN).

An assurance argument for an AI system certainly has to address the specific challenges posed by AI. The first step is to consider the safety requirements determined for the AI system, the AI's relationship with the overall system, and the input space. This input space describes the AI system's possible input values and is related to the Operational Design Domain (ODD). An AI system can only be audited for potential functional insufficiencies within the open-world context if the ODD specification goes into adequate detail [5]. The AI system's ability to operate in the input space has to be demonstrated qualitatively and quantitatively. This requires systematic analyses as well as statistical assessments, for example, to determine the probability of events occurring. The greater the input space's dimensions, the harder it is to achieve sufficient coverage. Engineers have to choose the right approach for each project. Not only the many standard scenarios have to be demonstrably 'covered, but also rare but critical situations. And the probability that unknown conditions could trigger functional deficiencies has to be low (see ISO 21448).

AI SAFETY LIFECYCLE

Engineers tasked to develop an AI system have to model its safety-relevant activities in the context of the overall system. To this end, they create an AI safety lifecycle as set out in [3].

The example of an AI safety lifecycle depicted in **FIGURE 1** shows all tasks necessary to help assert and sustain the assurance argument. As a rule, engineers deduce the safety requirements for the AI system from the encompassing system requirements. These requirements provide the input for developing, verifying, and validating the AI system. Datasets used to develop ML-based AI systems are usually generated on the fly as the AI component's iterative implementation progresses. Often, the right algorithm and methods for the given AI requirements do not begin to emerge until development is underway. This is why continuous and iterative safety analyses are imperative when developing AI systems. Usually, it is not possible to guarantee that an up-and-running AI system will always behave correctly. This is down to the ODD, which is highly complex with a status that varies over time, and to the nature of AI systems. And that is why an AI system has to be monitored and audited continuously. This process of substantiating the assurance argument's perpetual validity is called continuous assurance. If the argument is found to be invalid, engineers have to take remedial action by adapting the AI system, changing its requirements, or taking measures at a higher system level to re-establish its validity.

DATA LIFECYCLE

The behavior of ML-based AI systems is largely determined by the properties of the data used during development. These change during development and throughout the product lifecycle as new knowledge emerges and the environment changes. Requirements for a data lifecycle are a prudent way of assuring that current and consistent data is used to develop the AI. [3] Then the data

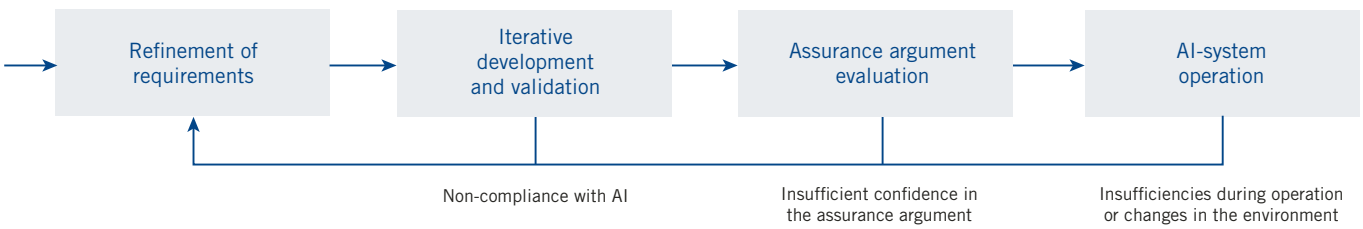


FIGURE 1 An example of an AI safety lifecycle – modified representation in line with [3] (© ITK Engineering GmbH)

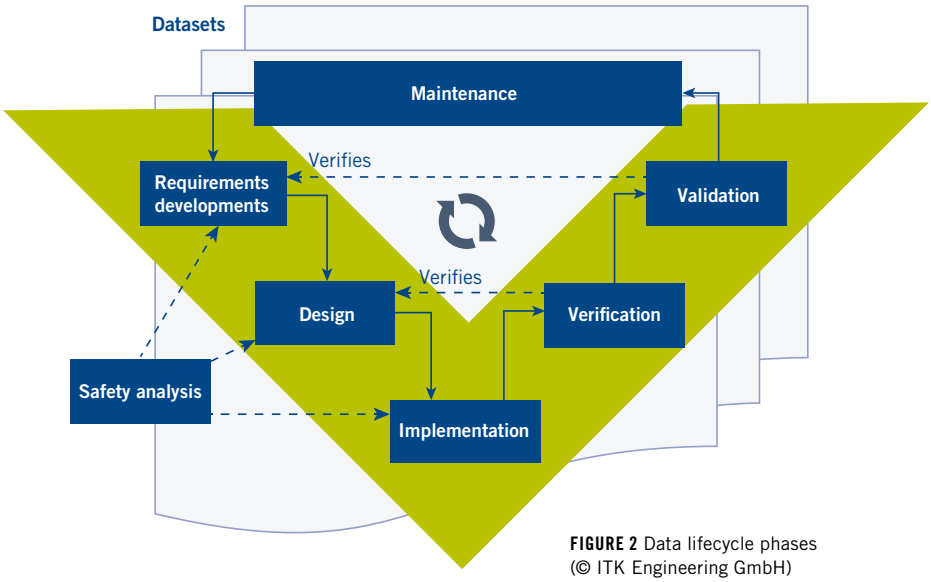


FIGURE 2 Data lifecycle phases (© ITK Engineering GmbH)

lifecycle can provide a basis for constructing a valid assurance argument.

The data lifecycle depicted in **FIGURE 2** typically begins with a data safety analysis. This aims to identify potential safety-relevant deficiencies, develop countermeasures, and define metrics for assessing preventive actions. The identified insufficiencies, root causes, and effects serve as the input for the subsequent phases, dataset requirements development, dataset design, and dataset implementation.

Dataset requirements development is all about formulating the specifications for the dataset and subsequently deducing the quality assurance requirements. This is done under the assumption that the methods set out in ISO 26262-3 for the item definition and ISO 21448 section 7 for the triggering conditions have been followed.

These requirements furnish the foundation for the next phase, dataset design, which revolves around the task of building the dataset. This entails compiling data from physical, synthetic, [6] or augmented data, preprocessing data, and dealing with metadata. Engineers then shall make their preparations, define annotation processes and methods, and carry out the actual annotation during the dataset implementation phase. Next up is dataset verification, which goes to ensure datasets are consistent, correct, and in line with the given requirements. Dataset validation serves to confirm that datasets satisfy the extrapolated requirements and meet expectations. Maintenance means keeping the working datasets up to date and in compliance.

DETERMINING AI REQUIREMENTS

Engineers are supposed to deduce the requirements for an AI system from the overall system's requirements, but these specifications are usually not quite specific enough to define direct measures to be taken in case of noncompliance.

To minimize risks during the AI development effort, it is advisable to define the influencing factors that will enable engineers to set out qualitative requirements. One way to get started is to consider the safety concerns mentioned in [7].

Some of these are measurable properties of the trained AI model, the data, or the development processes. Robustness, generalizability, explainability, and the like can be quantified, so they provide the metrics for refining AI safety requirements.

This is done by conducting a safety analysis to identify and assess properties that are maximally correlated with the violation of an AI requirement. Many of these properties can be determined directly or by way of correlated measurements. It is up to engineers to decide which of the various methods available for the given application and model are best suited for the use case. This assessment provides the insights they need to come up with specific measures to address the development or architecture of the AI system.

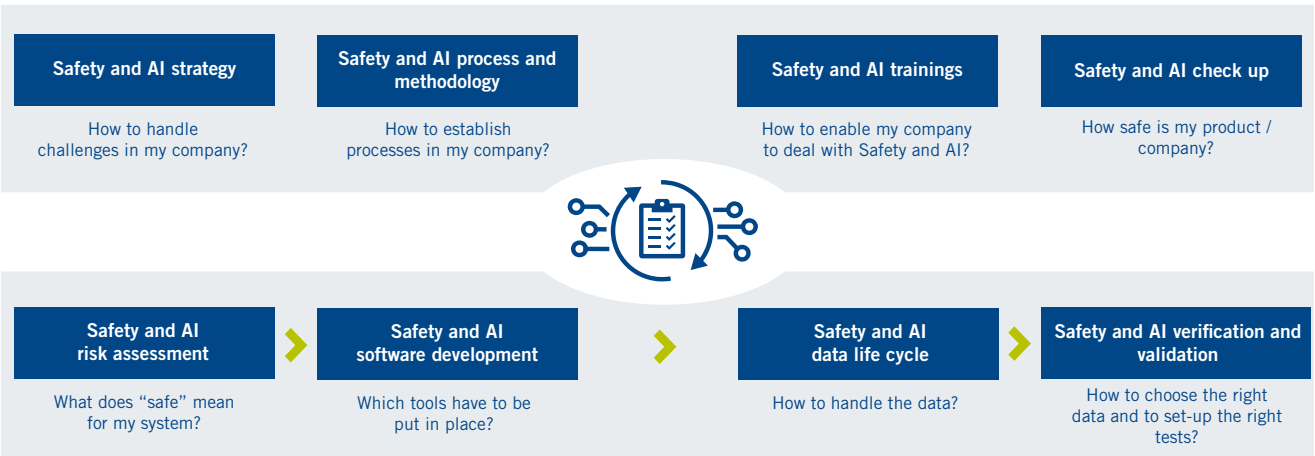


FIGURE 3 Safety and AI-specific challenges (© ITK Engineering GmbH)

CONCLUSION AND OUTLOOK

Emerging standards and specifications such as ISO PAS 8800 will certainly have an impact on the state of the art in automotive safety and AI. Engineers will need to meet certain requirements and follow specific processes for AI and data lifecycles to accommodate changes in safety requirements over time and after the start of production.

Data is integral to ML systems' training. It is even replacing some of the specifications and requirements, prompting the observation that data are the new requirements. This is why a systematic approach in form of a data lifecycle is sure to figure prominently in the future.Engineers have to get a holistic picture of safety aspects and factor the specificities of AI methods into the equation. To this end, they need additional safety analytics such as GSN and STPA.

Future safety and AI audits and assessments are going to pose a major challenge. Safety experts will have to add AI to their skill-sets – otherwise, outside AI experts will have to be brought in to provide support.

What's more, continuous development and continuous assurance represent a paradigm shift in safety engineering, so there is little experience to fall back on. In this respect, AI is both a blessing and a curse for safety-related automotive applications. On the one hand, AI can provide powerful solutions to problems such as environment recognition. On the other, developing and operating safe AI systems is a journey fraught with towering challenges.

A deep well of knowledge of specialists in safety, AI, and verification and validation creates precisely the skill-set needed to rise to safety and AI-specific challenges and solve emerging problems, **FIGURE 3**.

REFERENCES

[1] National Transportation Safety Board: Automated Vehicles – Investigations. Online: https://www.ntsb.gov/Advocacy/safety-topics/Pages/automated-vehicles-investigations.aspx, access: May 15, 2024

[2] EUR-Lex: Document 52021PC0206. Online: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52021PC0206, access: May 15, 2024

[3] Burton, S.: Standardisation of safe, data-driven AI Development & Tooling. KI Data Tooling Final Event. Online: https://www.ki-datatooling.de/fileadmin/KI_DataTooling/Downloads/Final_Results/FE_Presentations/20231205_KI_Data_Burton.pdf, access: May 15, 2024

[4] Klaes, M. et al.: Using Complementary Risk Acceptance Criteria to Structure Assurance Cases for Safety-Critical AI Components. Proceedings of the Workshop on Artificial Intelligence Safety 2021, Vol-2916, Online: https://ceur-ws.org/Vol-2916/paper_9.pdf, access: May 15, 2024

[5] ISO 34503: Road Vehicles – Test scenarios fo automated driving systems – Specification for operational design domain

[6] ITK Engineering GmbH: Individual Virtual Environment and Sensor Simulation (iVESS). Online: https://www.itk-engineering.de/en/story/individual-virtual-environment-and-sensor-simulation-ivess/, access: May 15, 2024

[7] Abrecht, S. et al.: Deep Learning Safety Concerns in Automated Driving Perception. Online: https://arxiv.org/pdf/2309.03774v1, access: May 15, 2024

THANKS

The authors would like to thank Philipp Leopold, AI expert at ITK Engineering, for his support in writing this article.

IMPRINT:
Special Edition 2023 in cooperation with ITK Engineering GmbH, Bergfeldstraße 2, 83607 Holzkirchen; Springer Fachmedien Wiesbaden GmbH, Postfach 1546, 65173 Wiesbaden, Amtsgericht Wiesbaden, HRB 9754, USt-IdNr. DE81148419

MANAGING DIRECTORS:
Stefanie Burgmaier I Andreas Funk I Joachim Krieger
PROJECT MANAGEMENT: Anja Trabusch

COVER PHOTO: © istockphoto I metamorworks

ITK Engineering

Stability, reliability and methodological expertise – this is what we have stood for since our founding in 1994. At all times, our customers have benefitted from our dedicated multi-industry know-how, especially in the fields of control systems design and model-based design. Customers can count on us – from conception through to deployment, we cover the entire development process.

Our areas of expertise include:

- Software development
- Hardware development
- Electrical & electronic systems
- System integration
- Software as a product
- Turnkey systems
- Customer specific development
- Technical consulting
- Seminars
- Quality assurance

The satisfaction of each of our partners and mutually respectful cooperation shape our corporate philosophy, in which four values are firmly anchored: Read more about this on the web.

V.1.0.0_e_2021

ITK Engineering GmbH
Headquarters: Ruelzheim
Im Speyerer Tal 6
76761 Ruelzheim, Germany
T: + 49 (0)7272 7703-0
F: + 49 (0)7272 7703-100
info@itk-engineering.com

www.itk-engineering.com

Founded in 1994
Branch offices throughout Germany – ITK companies worldwide.

Follow us on: