

ATZ extra



DATENGETRIEBENE ENTWICKLUNG
für Emotions- und
Müdigkeitserkennung nutzen

itk
ENGINEERING



© shutterstock | temp-64GTX | depositphotos | kerenby

Datengetriebene Entwicklung in der Automobilindustrie

VERFASST VON



Marcus Weber

ist Gruppenleiter Data Driven
Software and Sensors bei
ITK Engineering in Rülzheim.



Stefan Held

ist Lead Engineer bei
ITK Engineering in Holzkirchen.

Daten spielen für ein funktionierendes Mobility-Ökosystem eine Schlüsselrolle. Sie werden im laufenden Betrieb der Fahrzeuge genutzt, genauso wie bereits in der Entwicklung. Basis dafür sind datengetriebene Entwicklungsmethoden. Wie sie funktionieren, zeigt ITK Engineering am Beispiel der Emotions- und Müdigkeitserkennung, wie sie ab 2024 verpflichtend sein wird.

■ Daten, die eine Organisation generiert und besitzt, gelten zunehmend als wichtiger Unternehmenswert [1]. Damit einher geht die Hoffnung, dass sich die heute anfallenden Aufwände, um diese Daten nutzbar zu machen, künftig in einem Mehrwert für Unternehmen widerspiegeln. Während man allerdings beim Aufbau traditioneller Vermögenswerte langfristige Strategien entwickelt,

herrschen in der Datenökonomie weitläufig kurzfristige Perspektiven vor. In der Folge werden anwendungsbezogene Initiativen mit isolierten Datenszenen gestartet, anstatt eine skalierbare, unternehmensweit nutzbare Datenarchitektur aufzubauen. Zusätzliche Aufgaben und Verantwortlichkeiten werden bestehenden Rollen und Abteilungen zugewiesen, anstatt neue Mitarbeiterprofile zu entwi-

ckeln. Signifikante Datennutzungspotenziale werden aufgrund der fehlenden Integration der Initiativen in die Gesamtorganisation vernachlässigt.

Mit einer umfassenden Datenstrategie können diese Fehler frühzeitig vermieden werden. Sie bildet den Rahmen, Initiativen einzubetten und stellt sicher, dass aus Daten der größtmögliche Mehrwert generiert werden kann. Bei der Entwicklung der Strategie hilft ein agiles Vorgehen, um Kursänderungen zu ermöglichen und das unternehmerische Risiko zu reduzieren. Darüber hinaus empfiehlt es sich, Daten frühzeitig in die Entscheidungsfindung einzubeziehen. **BILD 1** zeigt eine beispielhafte Vorgehensweise, mit der datenbasierte Innovationsideen und das Datenpotenzial schnell bewertet werden können. Gleichzeitig wird dabei das Fundament für eine langfristige Strategie gelegt.

Mit diesem Vorgehen kann der Use Case zur Müdigkeits- und Emotionserkennung aus technischer, wirtschaftlicher und nutzerorientierter Sicht beleuchtet werden. Verschiedene Lösungsoptionen zur technischen Umsetzung wie der Nutzung von Kameradaten, dem Lenkradeinschlagswinkel oder weiterer Sensoren zur Aufzeichnung von

Vitalfunktionen lassen sich anhand verfügbarer Daten explorativ hinsichtlich ihrer Machbarkeit evaluieren. Daraus können Maßnahmen abgeleitet werden als Basis einer ganzheitlichen Strategie. Dazu zählen neben technischen Anpassungen auch Data-Governance-Maßnahmen, die im Zuge regulatorischer Anforderungen wie dem Data & AI Act der EU zunehmend an Bedeutung gewinnen [2, 3].

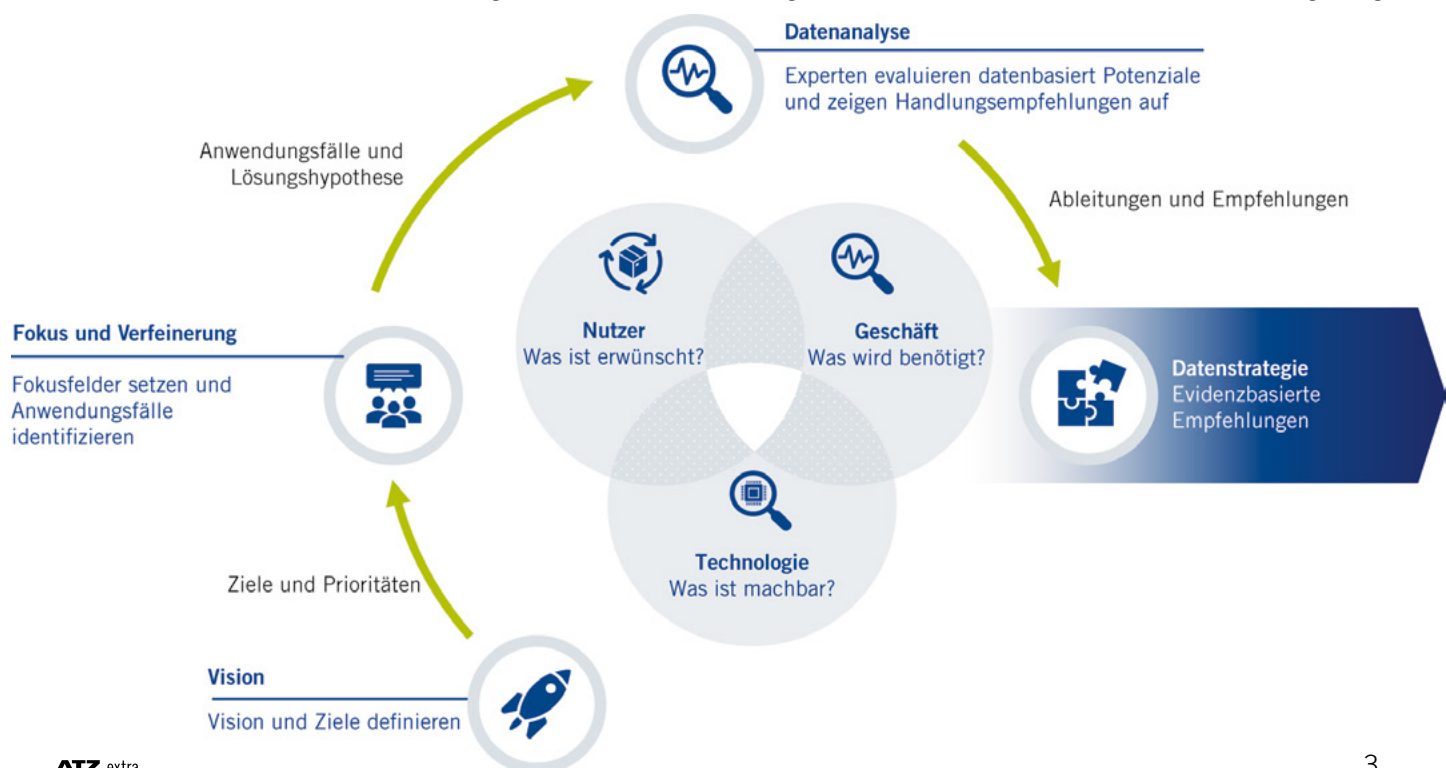
DATENQUALITÄT ALS VORAUSSSETZUNG FÜR PRÄZISE ANALYSEN

Die Datenqualität ist entscheidend, um Daten erfolgreich einzusetzen und genaue und zuverlässige Ergebnisse zu erzielen. Ihr wesentlicher Faktor ist die Eignung für die jeweilige Aufgabe. Um die Datenqualität zu messen und zu verbessern, müssen mehrere Faktoren berücksichtigt werden [4]:

- Die Daten sollten den spezifischen Anforderungen und Zielen des Entwicklungsprojekts entsprechen und müssen repräsentativ sein für alle Anwendungsfälle. Für den vorliegenden Use Case sind das etwaige Accessoires (Brillen, Kopfbedeckungen) als auch die Variation der menschlichen Erscheinung.

- Es sollten Nachweise erbracht werden, um zu zeigen, dass Datenfehler minimal sind. Auch Datenkorruption während der Speicherung oder Übertragung muss verhindert werden.
- Genauigkeit ist ein entscheidender Aspekt der Qualität, der insbesondere beim Labeling für ML-Entwicklung berücksichtigt werden muss.
- Auch die Vollständigkeit ist entscheidend. Durch die Festlegung von Anforderungen an den minimal akzeptablen Satz von Daten wird sichergestellt, dass alle relevanten Elemente des Eingabebereichs angemessen repräsentiert sind. Die Daten sollten der Variabilität der durch Szenarien beschriebenen Operational Design Domain [5, 6] entsprechen. Im vorliegenden Use Case reicht die Spannweite von unterschiedlichen Kameras und deren Positionen über Innenausstattung und deren Einfluss auf Sitzposition und Beleuchtung bis zu den Insassen.
- Rückverfolgbarkeit ist wichtig, um die Zuverlässigkeit und Glaubwürdigkeit der Daten beurteilen zu können.
- Data-Governance-Praktiken sind entscheidend, um hochwertige Daten zu erhalten. Dafür müssen Datenstandards und Validierungstechniken festgelegt sowie Qualitätskontrollprozesse durchgesetzt werden.

BILD 1 Eine umfassende Datenstrategie ist für Unternehmen die Grundlage, um aus vorhandenen Daten Mehrwert zu erzielen (© ITK Engineering)



DATENAUSWAHL UND ÜBERTRAGUNG IN DIE CLOUD

Ein zentraler Aspekt der datengetriebenen Entwicklung ist die Auswahl von Daten und die Übertragung vom Fahrzeug in die Cloud, wo sie verarbeitet, analysiert und genutzt werden, um verschiedene Funktionen im Automobilbereich zu entwickeln. Obwohl beispielsweise ML-Systeme mit Datensätzen trainiert werden könnten, benötigen sie oft zumindest Daten für das Transferlernen, um sich an spezifische Anwendungsfälle anzupassen. Der Use Case Müdigkeits- und Emotionserkennung umfasst spezielle Kamerapositionen, besondere Kameraphysik und die spezifischen Platz- und Lichtverhältnisse der Fahrzeuge. Ein wichtiger Schritt ist die Datenerfassung. Eine ungerichtete, dauerhafte Erfassung führt zu großen Datenmengen, die hohe Labelingaufwände verursachen. Alternativ ermöglicht die Festlegung von Kampagnen mit festen Grenzen die gezielte Erfassung von Daten innerhalb vordefinierter Parameter. Dieser Ansatz kann die Verwendung externer Trigger umfassen, um die Erfassung zu starten und zu stoppen. Dabei werden spezifische Analysetechniken angewandt wie beispielsweise der Vergleich mit spezifischen Szenarien oder beim Anwendungsfall der Müdigkeits- und Emotionserkennung stark springende Keypoints sowie räumlich zueinander versetzte Keypoints, die auf Probleme bei der Erfassung hinweisen könnten.

Die Vorteile der Übertragung von Daten in die Cloud sind zahlreich, insbesondere in Bezug auf die Data Gover-

nance, die durch ein zentralisiertes Gate in der Cloud gewährleistet werden kann. Darüber hinaus ermöglicht das Hochladen von Daten deren automatische Validierung und die Einrichtung spezifischer Pipelines für die Vorverarbeitung, Speicherung und Visualisierung. Zudem stehen sie unmittelbar zur Verfügung, was Arbeitsprozesse beschleunigt sowie Echtzeitanalysen und die Livepräsentation ermöglicht. Bei der Datenübertragung vom Fahrzeug in die Cloud muss Einiges beachtet werden: Der Zugriff auf den Cloud-Endpunkt muss gesichert sein, etwa durch Authentifizierung mittels eines Bearer-Tokens. Die Daten können innerhalb des Fahrzeugs zwischengespeichert und batchweise hochgeladen werden, was mobil oder stationär erfolgen kann. So können Keypoints [7] aufgrund der eher geringen Datenmenge mobil übertragen werden, Kameradaten werden zwischengespeichert und an Downloadstationen übertragen. Gesammelte Daten können in einem Eventstore wie Kafka für Persistenz und Nachverarbeitung gespeichert werden.

ZENTRALER DATENZUGRIFF FÜR DEZENTRALE DATENBASIS

Datengetriebene Entwicklungsansätze erfordern ein Ökosystem, das Entwickelnde befähigt, intuitiv Daten zu integrieren. Dafür muss ein niederschwelliger Zugang sichergestellt sein, genauso wie verlässlicher Inhalt und Rückverfolgbarkeit der eingesetzten Daten. Die Projektrealität zeigt, dass diese meist in verteilten, heterogenen Silos gespeichert werden. Der Zugriff ist in der Regel nicht

standardisiert und die nötigen Berechtigungen fehlen. Meist sind aufwendige und intransparente Genehmigungsworkflows der Grund. Auch der Auswertung selbst stehen oftmals komplexe Formate und die mangelnde Qualität der Daten im Wege.

Eine Lösung hierfür ist ein Digital Twin, der eine Abstraktionsebene zwischen Datenkonsumenten und Datensensoren bildet. Dabei fungiert er wie ein Wissensgraph als Repräsentation des realen Fahrzeugs und seiner Subsysteme. So ist es möglich, lesend und schreibend auf die Daten im Backend zuzugreifen, ohne ihre Spezifika zu kennen. Auch die Verifizierung der Daten und der Zugriffsberechtigungen können so abgebildet werden, **BILD 2**. Im betrachteten Beispiel können sich Entwickelnde über den Digital Twin einen robusten Datensatz zum Training des Modells zusammenstellen. Eine breite Nutzergruppe kann so verfügbare Daten kurzfristig erhalten, ohne dass Modifikationen an Datenarchitektur und Infrastruktur nötig sind. Durch das Aufbrechen der Datensilos entstehen zusätzliche Mehrwerte, die die Entwicklung neuer Services ermöglichen. Dieser Ansatz, kombiniert mit großen KI-Sprachmodellen, wird den Zugriff weiter erleichtern und die Datendemokratisierung beschleunigen [8].

VOM PROTOTYP ZUR SERIENENTWICKLUNG

Die Automatisierung des Lebenszyklus von ML-Modellen, von der Integration

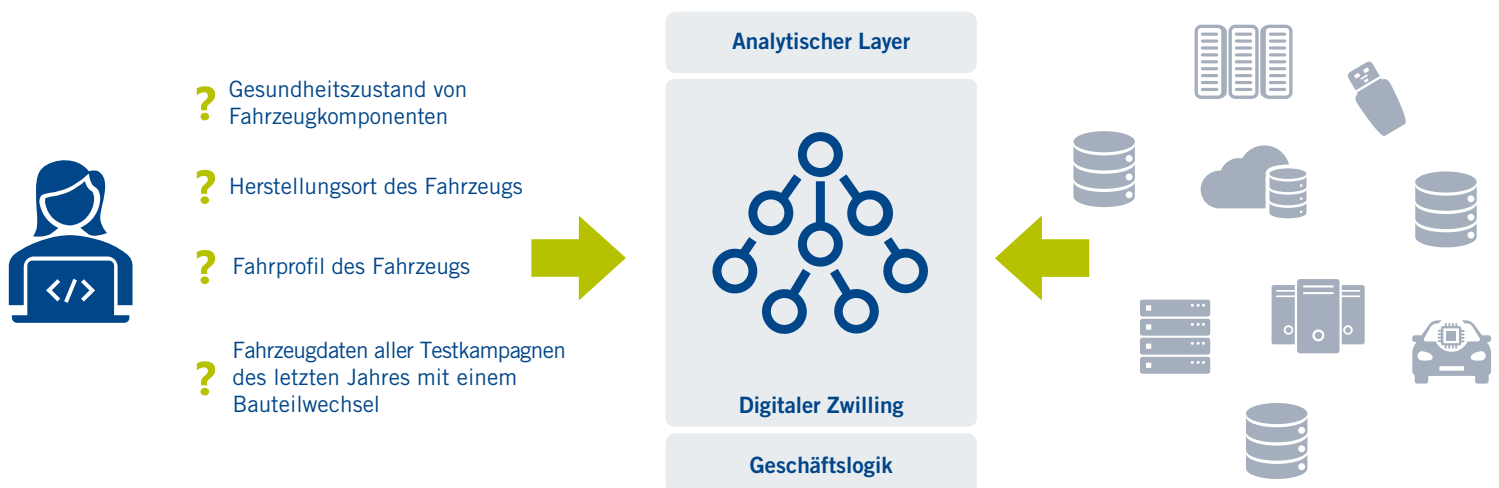


BILD 2 Ein Digital Twin kann Unternehmen helfen, das Maximale aus vorhandenen Daten herauszuholen (© ITK Engineering)

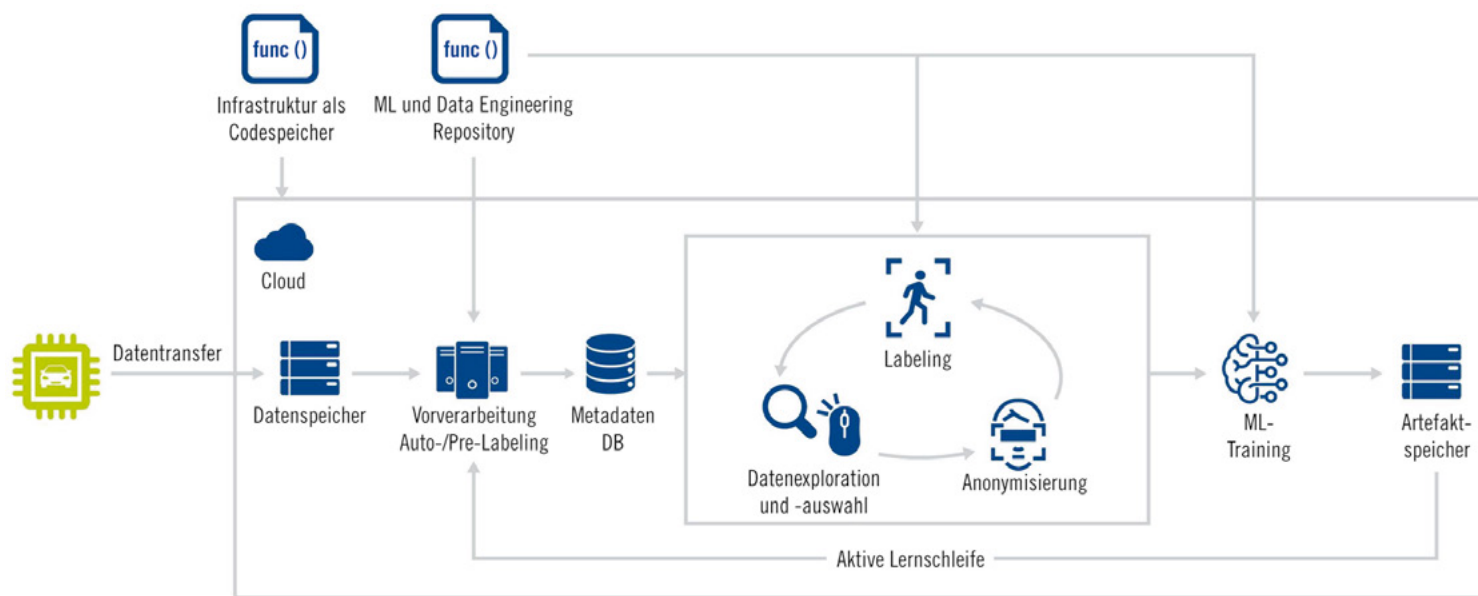


BILD 3 Ein lückenloser MLOps-Zyklus ist Voraussetzung für eine erfolgreiche Integration des ML-Modells in das Produkktivsystem (© ITK Engineering)

über die Bereitstellung bis zur Überwachung, wird als Machine Learning Operations (MLOps) bezeichnet [9, 10]. Bei der Entwicklung eines ML-Systems ist eine frühzeitig lückenlose Verarbeitungskette von der Quelle bis zur Datennutzung entscheidend, um die Kompatibilität der beteiligten Systeme sicherzustellen. Parallel zur Entwicklung dieser Pipeline sollten Daten-Labeling und Qualitätsverbesserung durchgeführt werden. Die Entwicklung von ML-Systemen ist ein iterativer Prozess aus den Teilschritten Modellerstellung, Test und Verfeinerung. Sobald die End-to-End-Pipeline einsatzbereit ist, folgt die Verbesserung der Genauigkeit des ML-Systems und die Integration von Methoden für erklärbare und verantwortungsvolle Explainable AI (XAI), um verständliche Erklärungen für die Entscheidungen von ML-Systemen zu liefern, Vertrauen aufzubauen und sicherzustellen, dass Vorurteile oder Fehler im Entscheidungsprozess erkannt und behoben werden. Im Use Case Müdigkeits- und Emotionserkennung sind Heat Maps sehr hilfreich.

Eine entscheidende Bedeutung bei der datengetriebenen Entwicklung spielt Skalierbarkeit. Ein System kann entweder vertikal skaliert werden, um die Kapazität einer Ressource zu erhöhen, oder horizontal durch Hinzu-

fügen von Rechnern oder Knoten. Beides kann die Kosten für die Bereitstellung und Wartung des Systems senken und das Training komplexerer Systeme ermöglichen. Der iterative ML-Entwicklungsprozess wird dabei mittels Leistungsmetriken des ML-Systems kontinuierlich überwacht. So kann das System auch dann optimal funktionieren, wenn sich Daten und Anforderungen ändern (Daten- und Konzeptdrift), **BILD 3**.

ZUSAMMENFASSUNG

Datengetriebene Entwicklungsansätze erfordern eine passende Datenstrategie, damit die richtigen Ziele verfolgt, passenden Maßnahmen ergriffen und optimale Ressourcen eingesetzt werden. So kann aus verfügbaren Daten mit der nötigen Qualität Mehrwert generiert werden. Daneben ist ein Ökosystem erforderlich, das Entwickelnde befähigt, intuitiv Daten zu integrieren. Ein Digital Twin reduziert Nutzungshürden und erzeugt eine digitale Repräsentation des Fahrzeugs und seiner Subsysteme. Weitere Erfolgsfaktoren sind die passende Projektstrategie und speziell für ML-Entwicklung eine unterstützende MLOps-Pipeline. Letztendlich erfordert die datengetriebene

Entwicklung in der Automobilindustrie technisches Know-how, Erfahrung und Leidenschaft.

LITERATURHINWEISE

- [1] Lempp, M.; Siegfried, P. (Hg.): Automotive disruption and the Urban Mobility Revolution. Cham: Springer, 2022, S. 35-37
- [2] Europäische Kommission (Hg.): Gestaltung der digitalen Zukunft Europas – Datengesetz. Online: <https://digital-strategy.ec.europa.eu/de/policies/data-act>, aufgerufen: 30. Juni 2023
- [3] Future of Life Institute (Hg.): The AI Act. Online: <https://artificialintelligenceact.eu/>, aufgerufen: 30. Juni 2023
- [4] Safety-Critical Systems Club (Hg.): SCSC Publications. Working Group Documents. Online: <https://scsc.uk/scsc-127H?tab=wgdoc>, aufgerufen: 30. Juni 2023
- [5] Pegasus (Hg.): PEGASUS Sicherheitsargumentation. Online: <https://www.pegasusprojekt.de/files/tmpl/pdf/PEGASUS%20Sicherheitsargumentation.pdf>, aufgerufen: 30. Juni 2023
- [6] Blank, F.; Hüger, F.; Mock, M.; Stauner, T.: Methodik zur Absicherung von KI im Fahrzeug. In: ATZ 124 (2022), Nr. 7-8, S. 58 - 63
- [7] ITK Engineering (Hg.): Individual Virtual Environment & Sensor Simulation: Driver/Occupant Monitoring Simulation. Online: <https://www.youtube.com/watch?v=Un8nOLCmIUQ>, aufgerufen: 30. Juni 2023
- [8] Patil, D.; Mason, H. (Hg.): Data Driven. O.O: O'Reilly Media, 2015
- [9] Luber, S.: Was ist MLOps?. Online: <https://www.bigdata-insider.de/was-ist-mlops-a-1116284/>, aufgerufen: 30. Juni 2023
- [10] Nayak, Y.: A Gentle Introduction to Mlops. Online: <https://towardsdatascience.com/a-gentle-introduction-to-mlops-7d64a3e890ff>, aufgerufen: 30. Juni 2023

IMPRESSUM

Sonderausgabe 2023 in Kooperation mit ITK Engineering GmbH, Bergfeldstraße 2, 83607 Holzkirchen; Springer Fachmedien Wiesbaden GmbH, Postfach 1546, 65173 Wiesbaden, Amtsgericht Wiesbaden, HRB 9754, USt-IdNr. DE81148419

GESCHÄFTSFÜHRER:

Stefanie Burgmaier | Andreas Funk | Joachim Krieger

PROJEKTMANAGEMENT: Anja Trabusch

TITELBILD: © Limitless Visions | stock.adobe.com



ITK Engineering

Seit der Firmengründung 1994 stehen wir für Stabilität, Sicherheit und Methodenexpertise. Damals wie heute bildet branchenübergreifendes Spezialwissen insbesondere im Bereich der Regelungstechnik und der modellbasierten Entwicklung die Basis, um unsere Kunden von der Idee bis zur Serienproduktion durchgängig und partnerschaftlich zu begleiten.

Unsere Kompetenzen umfassen u.a.:

- Softwareentwicklung
- Hardwareentwicklung
- Elektrik/Elektronik
- Systemintegration
- Software als Produkt
- Komplettlösungen
- Auftragsentwicklung
- Technische Beratung
- Schulungen
- Qualitätssicherung

Die Zufriedenheit all unserer Partner und ein respektvolles Miteinander prägen unsere Unternehmensphilosophie, in der vier Werte fest verankert sind: Lesen Sie gerne mehr darüber im Web.



V1.0.0_d_2021



ITK Engineering GmbH
Hauptsitz Rülzheim
Im Speyerer Tal 6
76761 Rülzheim
Tel.: + 49 (0)7272 7703-0
Fax: + 49 (0)7272 7703-100
info@itk-engineering.de

www.itk-engineering.de
www.itk-karriere.de

Gegründet 1994 –
heute hat ITK deutschland-
weit Niederlassungen und
ist international vertreten.

Folgen Sie uns auch auf:

