

IOSB

Fraunhofer

visIT

[KI-Engineering]

www.iosb.fraunhofer.de

ISSN 1616-8240



Fraunhofer
IOSB

Impressum

visIT erscheint etwa vier Mal pro Jahr und informiert über ausgewählte Forschungsthemen des Fraunhofer IOSB. Für die Bestellung von Einzelheften, ein kostenloses visIT-Abo sowie für Adressänderungen und Abbestellungen schicken Sie bitte eine E-Mail an publikationen@iosb.fraunhofer.de

Herausgeber

Prof. Dr.-Ing. habil. Jürgen Beyerer

Redaktion

Ulrich Pontes

Layout

Anja Wollfarth

Druck

Stork Druckerei GmbH
76646 Bruchsal

Anschrift der Redaktion

Fraunhofer-Institut für Optronik,
Systemtechnik und Bildauswertung IOSB
Fraunhoferstr. 1
76131 Karlsruhe
Telefon +49 721 6091-300
Fax +49 721 6091-413
presse@iosb.fraunhofer.de

© Fraunhofer IOSB
Karlsruhe 2021

Ein Institut der Fraunhofer-Gesellschaft
zur Förderung der angewandten
Forschung e. V. München

22. Jahrgang
ISSN 1616-8240

Bildquellen

Titelbild: CC-KING
Seite 4 + 5: Wikipedia
Seite 11: Samdraft auf Pixabay

Alle anderen Abbildungen:
© Fraunhofer IOSB

Nachdruck, auch auszugsweise,
nur mit vollständiger Quellenangabe und
nach Rücksprache mit der Redaktion.
Belegexemplare werden erbeten.

INHALT

Essay

- Seite 4 **Allianz Industrie 4.0 Baden-Württemberg**
Lukas Schleicher

Themen

- Seite 6 **KI-Engineering – Definition und Herausforderungen**
Julius Pfrommer
- Seite 8 **PAISE® – Das Vorgehensmodell für KI-Engineering**
Constanze Hasterok
- Seite 10 **Toolkit zur Erzeugung von Erklärungen für komplexe KI-Verfahren im Engineering-Bereich**
Elisabeth Peinsipp-Byma, Nadia Burkart
- Seite 12 **KI-Engineering für die Überwachung von Dampf- und Gasturbinen**
Christian Kühnert
- Seite 13 **KI-Engineering für die smarte Baustoffproduktion**
Christian Weißenbacher
- Seite 14 **Maschinelles Lernen und Simulation für automatisierte Mobilitätssysteme**
Masoud Roschani, Jens Ziehn

Liebe Freunde des Fraunhofer IOSB,

KI befasst sich mit der Schaffung technischer Systeme, die automatisiert kognitive und intellektuelle Aufgaben lösen können. Die Leistungsfähigkeit heutiger Rechner ermöglicht es, dass KI-Verfahren in ausgewiesenen Szenarien, z.B. bei der Objekterkennung in Bildern und Videos, deutlich effektiver und effizienter sind als der Mensch, vor allem dann, wenn Zusammenhänge und Muster in riesigen Datenmengen gelernt werden sollen.

Aber können KI-Methoden auch verlässlich und planbar in technischen Systemen eingesetzt werden? Wie muss man schon beim Entwurf vorgehen, damit sich eine Ingenieurin oder ein Ingenieur auf zugesagte Leistungsparameter verlassen kann? Das Fraunhofer IOSB hat sich auf die Fahnen geschrieben, diese Fragen mit einer neuen, mehrere Kompetenzfelder überspannenden Disziplin, dem KI-Engineering, zu erforschen und zu bearbeiten. Dieses Heft widmet sich dieser Thematik.

Der Gastbeitrag von Lukas Schleicher vom VDMA bzw. der Allianz Industrie 4.0 Baden-Württemberg stellt die Perspektive des Mittelstands beim Einsatz von Industrie 4.0 und KI dar.

Julius Pfrommer definiert den Begriff und die Dimensionen des KI-Engineering und erläutert die methodischen Herausforderungen. Dr. Constanze Hasterok präsentiert PAISE®, unser neues Vorgehensmodell für KI-Engineering. Elisabeth Peinsipp-Byma und Nadia Burkart beschäftigen sich mit der Frage der Erklärbarkeit in komplexen KI-basierten Systemen, ein wesentlicher Baustein für KI-Engineering.

In den abschließenden Beiträgen stellen Christian Kühnert, Christian Weißenbacher, Masoud Roschani und Jens Ziehn Anwendungsbeispiele des KI-Engineering für die Domänen Produktion und Mobilität vor. Die Praxisbeispiele stammen aus unserem Kompetenzzentrum für KI-Engineering Karlsruhe CC-KING, das wir als Fraunhofer IOSB gemeinsam mit dem FZI Forschungszentrum Informatik und dem Karlsruher Institut für Technologie (KIT) ins Leben gerufen haben. Unternehmen und wirtschaftsnahe Organisationen sind über den Innovationsbeirat eingebunden. Wenn Sie mehr darüber wissen wollen, schauen Sie vorbei auf den CC-KING-Seiten unter <https://www.ki-engineering.eu>.

KI-Engineering – ein spannendes und zukunftsweisendes Thema! Wir wünschen viel Freude und Anregungen bei der Lektüre unseres KI-Engineering-Heftes.

Fraunhofer IOSB, im Dezember 2021



Prof. Dr.-Ing. habil. Jürgen Beyerer Dr.-Ing. Thomas Usländer

Editorial



Prof. Dr.-Ing. habil. Jürgen Beyerer



Dr.-Ing. Thomas Usländer

ALLIANZ INDUSTRIE 4.0 BADEN-WÜRTTEMBERG

Baden-Württemberg ist geprägt – von was eigentlich? Man denkt gleich an ganz Unterschiedliches und doch an dasselbe: Wir sind nicht nur das Land der Tüftler und Visionärinnen, sondern haben auch landschaftlich einiges zu bieten: Barocke Schlösser oder der mystische Schwarzwald locken Touristen aus aller Welt an. Romantische Industriedenkmäler im Schwarzwald zeugen ebenso von hervorragender Ingenieurskunst wie ganze Städte, die nach den Vorstellungen damaliger Stadtplaner vom Reißbrett aus entworfen wurden. Und zu guter Letzt denkt man an bodenständig und nachhaltig agierende Familienunternehmen.

SORGFALT UND VIELFALT

Doch was das Land wirklich ausmacht ist die Kombination aus all dem. Das Hightech-

Unternehmen aus dem Schwarzwald, von dessen Innovationen die Lieferketten ganzer Branchen abhängen, steht exemplarisch dafür. Diese Unternehmen verkörpern genau diese scheinbaren Gegensätze: Hightech, ländlicher Raum, kleine Unternehmen, nachhaltiges Denken, kurze Wege und Tradition.

HERAUSFORDERUNGEN DES MITTELSTANDS

Wenngleich das Ländle hervorragende Expertise im digitalen Bereich aufweist, ergeben sich durch die digitale Transformation für sie ganz besondere Herausforderungen. Beispielsweise wurde in Baden-Württemberg die erste deutsche Informatikfakultät an der technischen Hochschule etabliert, deren Alumni auch den heute größten



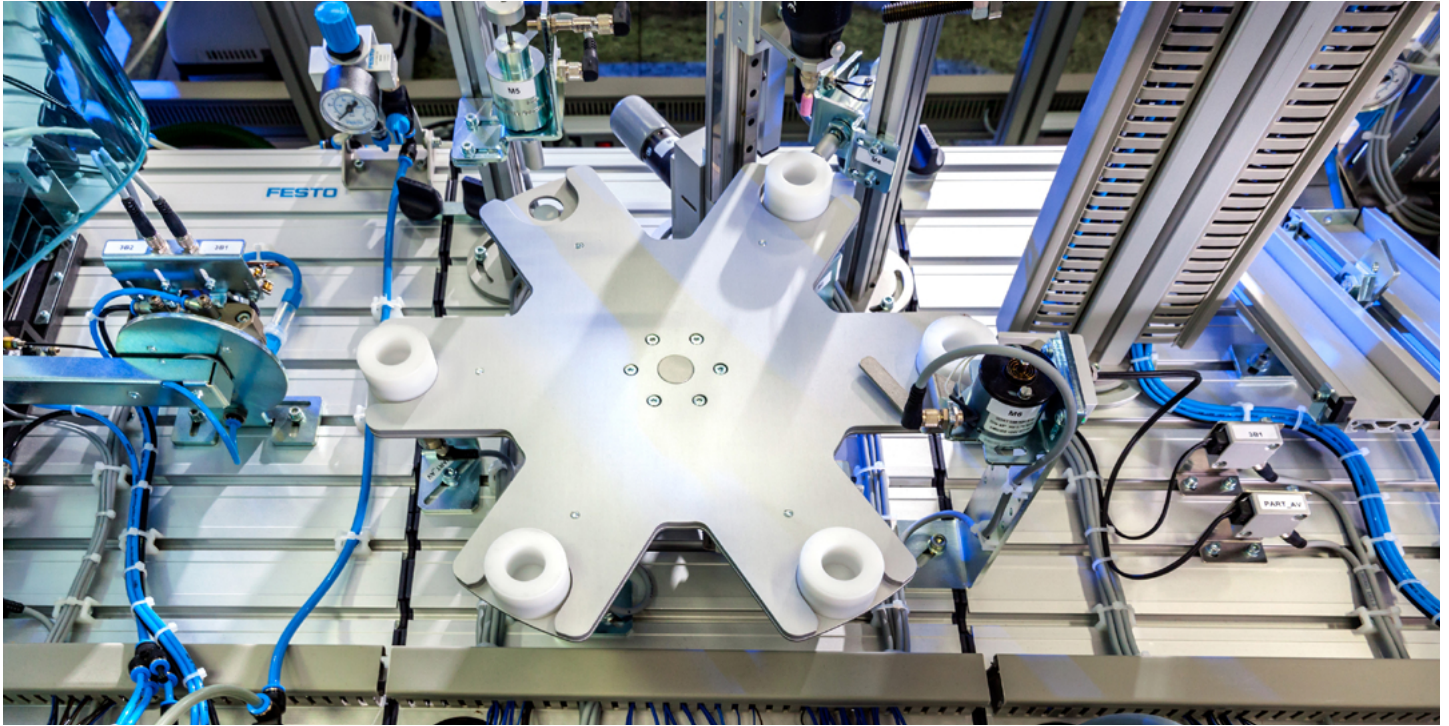
Lukas Schleicher M.Sc.

VDMA Baden-Württemberg/
Allianz Industrie 4.0
Baden-Württemberg

Telefon +49 711 22801 29
lukas.schleicher@vdma.org
<https://www.i40-bw.de>



Vom Patent-Motorwagen Nr. 3 bis zur automatisierten Fabrik hat sich vieles verändert, aber eines ist gleich geblieben: der Pioniergeist der Tüftler und die Innovationskraft des ganzen Landes.



Zur Zeit von Carl Benz hätte sich noch keiner etwas unter Automatisierungstechnik vorstellen können. Heute ist keine Fabrik mehr ohne möglich. Was mit Maschinellen Lernen zukünftig möglich ist, ist heute schwer abschätzbar.

europäischen Softwarekonzern gegründet haben. Um Fachkräfte buhlen nun etablierte Industriekonzerne, Silicon-Valley-Monopolisten und Familienunternehmen »von der Alb« gleichermaßen. Aber gerade bei digitalen Produkten kommt eine weitere Herausforderung hinzu: Viele Produkte skalieren mit der Größe. Die Implementierungskosten unterscheiden sich unwesentlich im Verhältnis zum Nutzen, der mit der Breite der Anwendung skaliert. Gerade bei kleinen Stückzahlen von hoch spezialisierten Anwendungen ergibt sich ein struktureller Nachteil: Die Sprachsteuerung eines Smartphones hat potenziell 100 Millionen (muttersprachliche) Anwender, womit das Reservoir an Datenspendern groß ist. Außerdem lassen sich Entwicklungskosten auch auf potenziell 100 Millionen Anwender umlegen. Die Zahl der Laserschneidemaschinen auf der Welt bewegt sich in einer wesentlich kleineren Größenordnung. Wenn man bei einer solchen Maschine die Prozesse datengetrieben optimieren will, ist das eine ganz andere Herausforderung: Die Datenmenge

ist viel geringer und die auf die einzelne Maschine umgelegten Kosten sind viel höher.

VORGEHENSMODELLE ETABLIEREN

Was also tun, um dem Problem zu begegnen? Es gibt viele Tech-Missionare und Evangelisten, Prediger der Transformation und Exorzisten des Etablierten. Sie reden aber an den tatsächlichen Bedarfen der Branche vorbei. Die Industrie ist sich den Chancen und Herausforderungen der digitalen Transformation durchaus bewusst. Die Allianz Industrie 4.0 Baden-Württemberg begleitet seit 2015 den Mittelstand bei den damit verbundenen Herausforderungen. Statt immer neuer Bergpredigten geht es vor allem darum, Netzwerke aus Anwendern zu schaffen. Konkrete Lösungsansätze und bereits realisierte Projekte helfen, die abstrakten Potenziale im eigenen Unternehmen zu erkennen!

Um »Künstliche Intelligenz (KI)« in die Breite zu bringen, kann auch KI-Engineering unter-

stützen. So komplex manches Datenmodell auch sein mag, Probleme liegen oft in organisatorischen Fragen. Dabei ergibt sich viel Potenzial, den Unternehmen die Umsetzung zu erleichtern: von der interdisziplinären Ausbildung und dem Etablieren einer eigenen Fachdisziplin eines KI-Ingenieurwesens bis hin zu standardisierten, systematischen Vorgehensweisen. Ziel soll sein, Projekte schnell und einfach skalieren zu können und früh Erfolgsaussichten des Projekts einzuschätzen!



Mehr Informationen zur Allianz Industrie 4.0 Baden-Württemberg unter www.i40.bw.de.

KI-ENGINEERING – DEFINITION UND HERAUSFORDERUNGEN

Systems Engineering ist die Disziplin, komplexe technische Systeme zu entwickeln und zu realisieren. Flugzeuge, Kraftwerke und Mobiltelefone sind Beispiele für komplexe Systeme, die von teils sehr großen Teams umgesetzt werden. Heute werden mehr und mehr KI-Lösungen in Gesamtsysteme integriert, um diese performanter, autonom handelnd oder sogar lernfähig und selbst-optimierend zu machen. Die Entwicklung KI-basierter Lösungen passiert aber noch vielfach ad-hoc und in einem kleinen Team – nicht unähnlich einem Forschungsprojekt. Im Kontrast dazu *adressiert KI-Engineering die systematische Entwicklung und den Betrieb von KI-basierten Lösungen als Teil von Systemen, die komplexe Aufgaben erfüllen*. Dabei wird das Kunstwort KI-Engineering auf Englisch übersetzt als »AI Systems Engineering«.

In der Folge werden die Herausforderungen von KI-Engineering anhand von vier komplementären Sichtweisen (siehe Abbildung 1) auf das Thema aufgezeigt.

ANWENDUNGSDIMENSIONEN

Die von KI-Engineering adressierten Anwendungen unterscheiden sich von denen »normaler KI«. Typischerweise treten eine oder mehrere der folgenden Eigenschaften zu Tage.

- **Kritikalität** bezieht sich auf die Auswirkungen eines nicht funktionierenden Systems auf Sicherheit (für Menschen und Systeme), geschäftskritische Funktionalität, Datenschutz und weitere Risiken. Wenn die Kritikalität hoch ist, werden spezielle Maßnahmen und möglicher-

weise eine offizielle Zertifizierung benötigt, um die korrekte Funktion des Gesamtsystems sicherzustellen, in das die KI-Lösung integriert ist.

- **Organisatorische Komplexität** bezieht sich auf den Mehraufwand, der nötig ist, um die Entwicklung und den Betrieb eines KI-Systems zu koordinieren. Dies ist besonders relevant bei Arbeiten in großen, heterogenen Teams oder bei notwendigen Abstimmungen, Austausch von Daten etc. über mehrere Firmen hinweg.
- **Physikalität** bezieht sich darauf, wie stark die Anwendung Bezug zur physischen Welt und eine direkte Beziehung zu den Naturwissenschaften (Physik, Chemie etc.) bzw. den traditionellen Ingenieursdisziplinen hat. Je unmittelbarer eine KI der physikalischen Welt zugehörig ist, desto mehr Vorwissen kann als »physikalische Grundwahrheit« integriert werden.

HERAUSFORDERUNGEN IM ENTWICKLUNGSPROZESS

Die Entwicklung KI-basierter Lösungen und insbesondere der Einsatz datengetriebener Verfahren ist häufig schwer mit klassischem Vorgehen in den Ingenieurdisziplinen vereinbar. Für die Entwicklung einer KI-Lösung werden meist iterativ (agil) Lösungsansätze umgesetzt und empirisch validiert. Das kann die Finalisierung des High-Level-Designs auch für die anderen Gewerke in einem Gesamtsystem verzögern. Kommen datengetriebene Verfahren (Machine Learning) zum Einsatz, so sind früh hinreichend Daten zu beschaffen, die denen aus dem Betrieb stark ähneln. Ist das Gesamtsystem aber erst in der Entwick-



Dr.-Ing. Julius Pfrommer

Informationsmanagement
und Leittechnik (ILT)
Fraunhofer IOSB Karlsruhe

Telefon +49 721 6091-286
julius.pfrommer@iosb.fraunhofer.de
www.iosb.fraunhofer.de/ILT

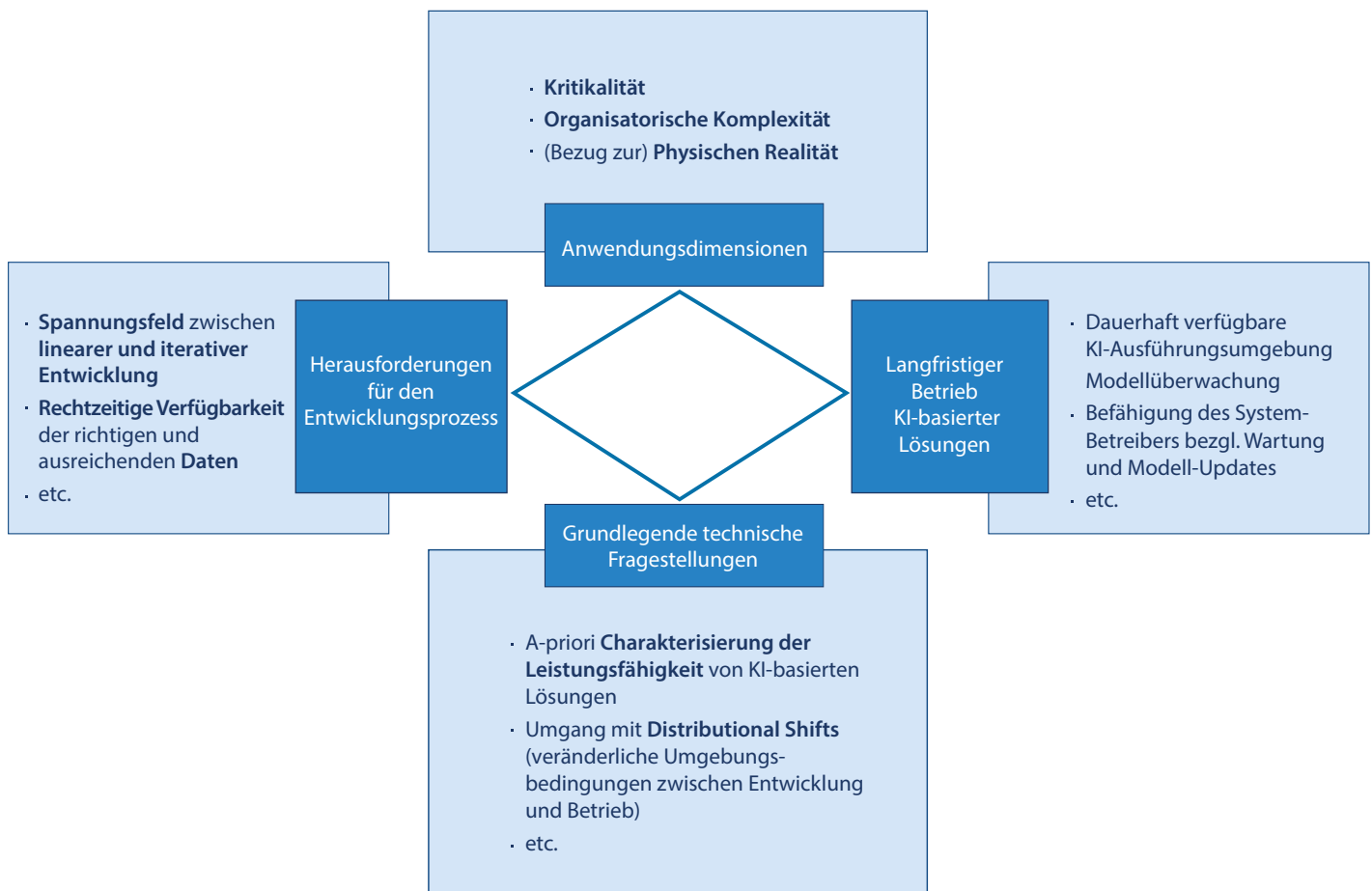


Abbildung 1: Strukturierung der Herausforderungen von KI-Engineering

lung, können Daten aus dem Betrieb manchmal noch nicht beschafft werden. Dann kann auf speziell entwickelte Simulations-Umgebungen zurückgegriffen werden.

LANGFRISTIGER BETRIEB KI-BASIERTER LÖSUNGEN

Komplexe technische Systeme sind häufig über Jahre und Jahrzehnte im Einsatz. In dieser Zeit entwickeln sie sich auch weiter. Über natürlichen Verschleiß und Wartung genauso wie über strukturelle Umbauten. Zunächst muss erkannt werden, ab wann die Annahmen der bisherigen KI-Lösung nicht mehr zutreffen und nachgebessert werden muss. Tritt der Fall ein, muss das System inklusive integrierter KI-Lösungen wieder in einen betriebsfähigen Zustand versetzt werden. Idealerweise wird der Betreiber befähigt Wartungen und kleine Modell-Updates selbst vorzunehmen.

Weiterhin ist sicherzustellen, dass die Ausführungsumgebung, in der die KI umgesetzt wird, über die gesamte Lebensdauer des Systems trägt.

GRUNDLEGENDE TECHNISCHE FRAGESTELLUNGEN

Obgleich KI-Engineering auf die Umsetzung komplexer Systeme im Unternehmenskontext hinarbeitet, wurden auch Themen identifiziert, die einen Beitrag der Grundlagenforschung erfordern. Zum einen sollte es für den systematischen Einsatz von KI möglich sein, deren Performanz früh im Entwicklungsprozess sicher abschätzen zu können. Heute kann in vielen Fällen die Performanz KI-basierter Lösungen nur empirisch und für einen bereits weit fortgeschrittenen Stand der Entwicklung gezeigt werden. Dies birgt Risiken für den Entwicklungsprozess. Eine weitere grundlegende Fragestellung

ist der fundierte Umgang mit sogenannten Distributional Shifts. Die Umgebung, aus der Datensätze für Maschinelles Lernen gewonnen werden, entspricht dann nicht der späteren Anwendungsumgebung. Ursache dafür ist etwa eine graduelle Veränderung der Anwendungsumgebung über die Zeit oder auch die Anwendung eines einmal trainierten Modelles in parallelen Instanzen der Anwendung, mit spezifischen (kleinen) Unterschieden.

KI-Engineering adressiert die genannten Herausforderungen mit dem Ziel eine Professionalisierung des Einsatzes von KI in komplexen Systemen zu erreichen.

Nähere Informationen zum Thema auf: <https://www.ki-engineering.eu/de/was-ist-ki-engineering.html>

PAISE® – DAS VORGEHENS-MODELL FÜR KI-ENGINEERING

PAISE® steht für **P**rocess Model for **AI** Systems **E**ngineering und bezeichnet das Vorgehensmodell für KI-Engineering. Es strukturiert den Entwicklungsprozess von KI-basierten Systemlösungen, wobei die Vorgehensweisen aus der Informatik und datengetriebenen Modellbildung mit denen klassischer Ingenieurdisziplinen kombiniert werden.

In den letzten Jahrzehnten rücken maschinelle Lernverfahren (ML-Verfahren) als Untergruppe von KI-Algorithmen durch die zunehmende Verfügbarkeit von Daten vermehrt in den praktischen Vordergrund. Der ML-Algorithmus programmiert einfach gesagt eine Software zur Erfüllung einer gegebenen Aufgabe, indem er sogenannte Trainingsdaten analysiert und dort Muster und Zusammenhänge identifiziert. Die Funktionen der erstellten Software werden daher zum wesentlichen Teil durch die Zusammensetzung und die Qualität der Trainingsdaten bestimmt. Dieser Ansatz bringt zwar einen immensen Effizienzgewinn in der Entwicklung mit sich, stellt jedoch folgende neue Herausforderungen an den Produktentwicklungsprozess: Zum einen lässt sich die Performanz einer KI-basierten Lösung, speziell beim Einsatz von maschinellen Lernverfahren, nicht im Voraus verlässlich einschätzen, sondern muss empirisch ermittelt werden. Zum anderen werden bereits während der Entwicklung die Daten aus dem Betrieb benötigt, die oftmals nicht vorhanden sind.

Das Vorgehensmodell PAISE® adressiert diese Herausforderungen und ist schematisch in Abbildung 1 dargestellt.

Im Sinne einer ingenieurmäßigen Vorgehensweise beleuchtet PAISE® die Entwicklung eines Produkts als ein Gesamtsystem, das in

Subsysteme mit klar definierten Schnittstellen zerlegbar ist. Diese Funktionale Dekomposition findet in der dritten Phase des Vorgehensmodells statt. Hier wird ein Systemmodell initial erstellt, welches als durchgehendes Artefakt während des Entwicklungszyklus in der fünften Phase kontinuierlich angepasst und verfeinert wird. ML-Verfahren können zum einen direkt in Subsysteme integriert werden, zum anderen können sie bei der Entwicklung von Hilfssystemen (engl. Enabling Systems) eingesetzt werden, die nicht Teil des Endprodukts sind. Soll zum Beispiel ein Notbremsssystem entwickelt werden, welches im Gefahrenfall ein Auto autonom abbremst, so kann dies durch Detektion von Objekten in Kamerabildern erfolgen. Zum Gesamtsystem gehören in diesem Fall die Kamera, ein Subsystem, und der Detektor, welcher als KI-basiertes Subsystem (KI-Komponente) entwickelt wird. Hilfssysteme sind in diesem Szenario z. B. die Software, welche für die gegebene Kamera eine Optimierung der Befestigungsposition durchführt oder eine Datenbank, in der Aufnahmen der Kamera gespeichert werden können.

Eine Besonderheit von PAISE® ist, dass auch Datensätze, welche für die Entwicklung von KI-basierten Subsystemen benötigt werden, im Systemmodell verortet werden. So wird nicht nur ein Vorgehen für die Entwicklung von ML-Komponenten vorgegeben, sondern auch für die Entwicklung von Datensätzen. Beide Vorgehen haben gemeinsam, dass sie gemäß eines V-Modells strukturiert sind und die Ergebnisse auf Basis von wohldefinierten Anforderungen und Metriken evaluiert werden.

Charakteristisch für PAISE® ist der Entwicklungszyklus in der fünften Phase, der an agile Methoden aus der Softwareentwick-



Dr. Constanze Hasterok

Informationsmanagement
und Leittechnik (ILT)
Fraunhofer IOSB Karlsruhe

Telefon +49 721 6091-602
constanze.hasterok@iosb.fraunhofer.de
www.iosb.fraunhofer.de/ILT

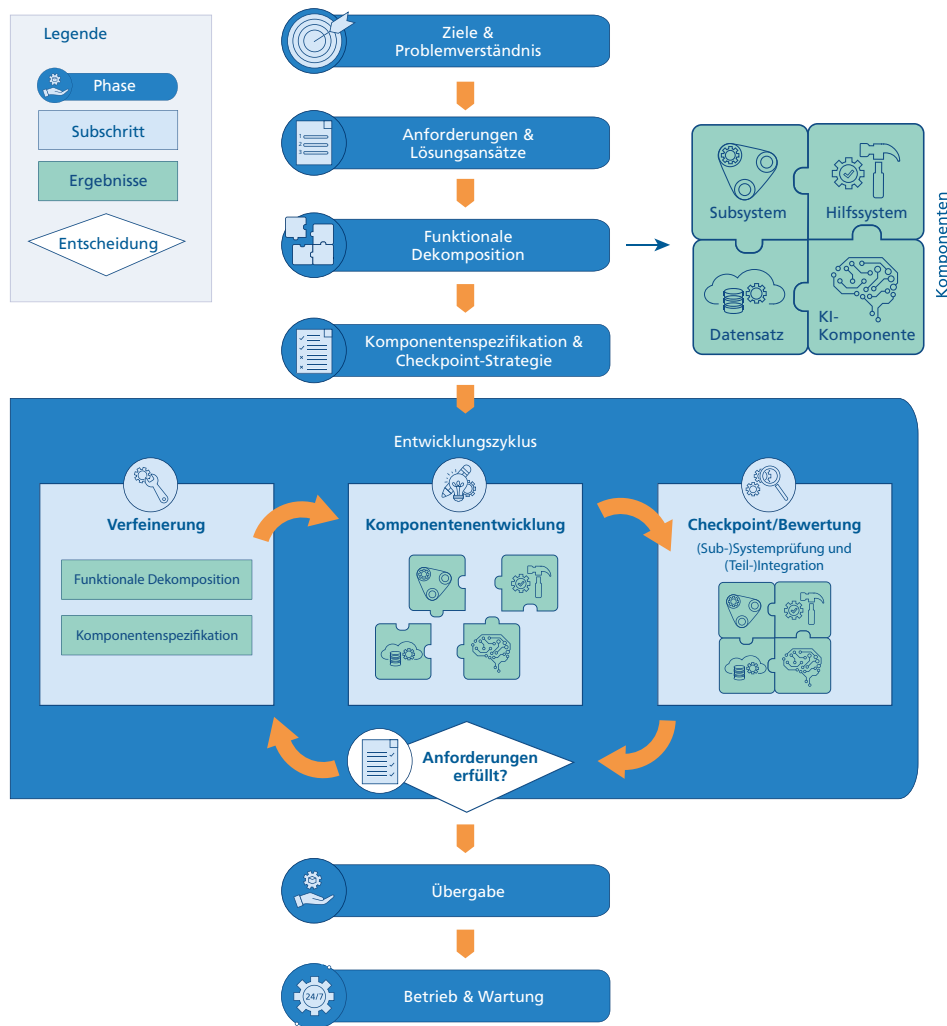


Abbildung 1: Schematische Darstellung des Vorgehensmodells PAISE®

lung angelehnt ist. Der Entwicklungszyklus besteht aus den Schritten Verfeinerung, Komponentenentwicklung und Checkpoint, deren iteratives Durchlaufen die empirische Ermittlung der Performanz von KI-basierten Komponenten erlaubt. Die Komponentenentwicklung findet für alle Komponenten parallel und unabhängig voneinander statt. Zu den Checkpoints werden die Komponenten in das Gesamtsystem (teil-)integriert und dieses wiederum hinsichtlich der Anforderungen bewertet. Am Beispiel des Notbremsystems kann ein exemplarischer Entwicklungszyklus folgendermaßen aussehen: In der Komponentenentwicklung wird der Detektor auf Basis erster frei verfügbarer Datensätze und initial abgeschätzten Spezifikationen entwickelt, während die Kamera zunächst ausgewählt und beschafft wird.

Beim Checkpoint werden die Schnittstellen der Kamera und des Detektors miteinander verknüpft und Live-Bilder der Kamera vom Detektor ausgewertet. Es stellt sich heraus, dass die Objekterkennung nicht gut funktioniert, vermutlich weil die Auflösung der Kamera eine andere ist als die der Bilderdaten, auf denen der Detektor initial entwickelt wurde. In der darauffolgenden Verfeinerung wird beschlossen, einen Trainingsdatensatz zu entwickeln unter Nutzung der vorhandenen Kamera. Gleichzeitig soll der Detektor robuster werden gegenüber unterschiedlichen Kameraauflösungen, sodass auch andere Kamerateypen verwendet werden können. In der folgenden Komponentenentwicklung wird der Detektor mit frei verfügbaren Datensätzen anderer Auflösung weiterentwickelt und der Datensatz nach

dem von PAISE® vorgegebenen Vorgehen entwickelt. Ist dieser vorhanden, findet eine Übergabe an die Detektorentwicklung zum nächsten Checkpoint statt. Im nächsten Zyklus kann nun der Detektor auf Basis des neuen Datensatzes weiterentwickelt werden.

Das Vorgehensmodell wurde innerhalb des Kompetenzzentrums für KI-Engineering (CC-KING) entwickelt und wird nun zusammen mit kleinen und mittelständischen Unternehmen erprobt und verfeinert.

Nähere Informationen zum Thema auf: <https://www.ki-engineering.eu/de/wissen-tools/paise.html>

Gefördert wird CC-KING vom Ministerium für Wirtschaft, Arbeit und Tourismus des Landes Baden-Württemberg.

Themen



Dr. Elisabeth Peinsipp-Byma

Interaktive Analyse und
Diagnose (IAD)
Fraunhofer IOSB Karlsruhe

Telefon +49 721 6091-393
elisabeth.peinsipp-byma@iosb.fraunhofer.de
www.iosb.fraunhofer.de/IAD



Nadia Burkart M.Sc.

Interaktive Analyse und
Diagnose (IAD)
Fraunhofer IOSB Karlsruhe

Telefon +49 721 6091-619
nadia.burkart@iosb.fraunhofer.de
www.iosb.fraunhofer.de/IAD

TOOLKIT ZUR ERZEUGUNG VON ERKLÄRUNGEN FÜR KOMPLEXE KI-VERFAHREN IM ENGINEERING- BEREICH

Künstliche Intelligenz (KI) spielt eine immer wichtigere Rolle, da sie große Datenmengen schnell analysieren und basierend darauf hilfreiche Hinweise generieren kann. So können KI-basierte Lösungen in der Produktion Ausfälle bestimmter Produktionsanlagen gezielt vorhersagen und damit die Möglichkeit geben, vorzeitig entgegenzuwirken. Die KI-Verfahren bzw. deren interne Funktionsweisen sind für den Menschen jedoch oftmals kaum bis gar nicht zu verstehen und werden daher als »Black-Box-Modelle« bezeichnet. Insbesondere die in den letzten Jahren immer erfolgreicher gewordenen tiefen neuronalen Netze gehören zu dieser Art von KI-Verfahren. Sie bieten zwar eine hohe Vorhersagequalität, liefern aber keine Erklärungen für die Vorhersagen. Eine zunehmend wichtige Voraussetzung zum Einsatz einer KI ist jedoch, dem Menschen die Funktionsweise der Vorhersagemodelle verständlich zu beschreiben. Nur so kann der Mensch dem KI-Verfahren und seinen Hinweisen vertrauen und entsprechend reagieren. Bedingt durch diesen Umstand rückt die Erklärbarkeit der Black-

Box-Modelle von KI-Verfahren immer stärker in den Vordergrund.

Das Forschungsfeld Explainable Artificial Intelligence (XAI) beschäftigt sich mit der Generierung nachvollziehbarer Erklärungen von Ergebnissen aus komplexen KI-Verfahren. Hierzu wird am Fraunhofer-Institut für Optoelektronik, Systemtechnik und Bildauswertung IOSB seit vielen Jahren daran geforscht, wie Erklärungen für Entscheidungen von Blackbox-Modellen erzeugt werden können und welche Erklärungen den Menschen beim Verständnis für die KI-Entscheidung bestmöglich unterstützen. Dazu wurde eine XAI-Toolbox konzipiert, die für verschiedene KI-Verfahren bzw. die dahinter stehenden Blackbox-Modelle eingesetzt werden kann. Die XAI-Toolbox erzeugt schnell und zielgerichtet Erklärungen für ein Blackbox-Modell und stellt dazu verschiedene vorimplementierte XAI-Methoden zur Verfügung. Dabei wird in der Toolbox zwischen zwei Arten von Erklärungen unterschieden: den globalen und den lokalen Erklärungen. Globale Erklärungen beschreiben das Verhalten des

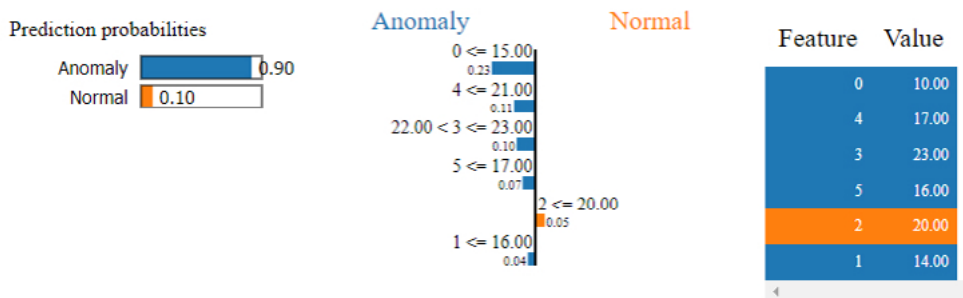


Abbildung 1: Lokale Erklärung für die Vorhersage einer Anomalie.

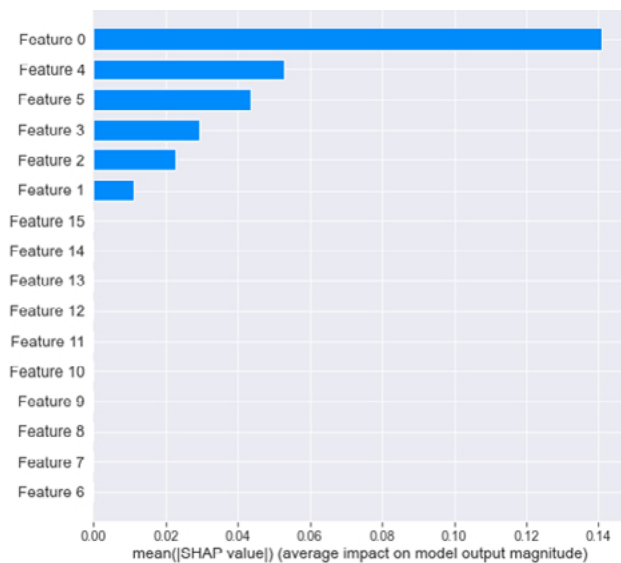


Abbildung 2: Globale Erklärung des Modells.

gesamten Modells. D. h. es werden Abhängigkeiten über den gesamten gelernten Datensatz dargestellt. Lokale Erklärungen beziehen sich auf das Vorhersageergebnis des Modells zu einem einzelnen Datensatz. Es werden die Abhängigkeiten in Bezug auf eine einzelne Vorhersage beschrieben. Die Toolbox kann zur Erklärung des gesamten Modells bzw. KI-Verfahrens sowie zur Erklärung einzelner Vorhersagen des KI-Verfahrens eingesetzt werden. Damit kann die Toolbox sowohl Entwickler und Entwicklerinnen des KI-Verfahrens als auch Nutzer und Nutzerinnen beim Einsatz des KI-Verfahrens durch das jeweils erforderliche Verständnis für das Verhalten der KI unterstützen.

Beispielsweise ist es im Bereich der Produktion und der Instandhaltung von Anlagen wichtig, frühzeitig Anomalien und dadurch eventuell verursachte Störungen zu erkennen. Dadurch können Stillzeiten und mögliche Produktionsengpässe verhindert werden. Wenn dann ein KI-Verfahren – welches zur Unterstützung bei der Instandhaltung einer Produktionsanlage entwickelt wurde – vorhersagt, dass eine bestimmte Störung

auftreten wird, ist es für die jeweiligen Mitarbeitenden wichtig zu wissen, wodurch diese auftreten. Hier helfen Erklärungen, gezielt die entsprechenden Schwachstellen der Anlagen zu lokalisieren. Zudem können durch die Erklärungen die Störungen ganz konkret behoben werden, da genau ersichtlich ist, welche Komponente (z. B. ein Ventil) für die Störung zuständig ist. In Abbildung 1 ist das Ergebnis einer lokalen Erklärung für die Vorhersage einer Anomalie dargestellt.

Darauf ist links oben zu erkennen, dass der Eintritt einer Anomalie mit 90 Prozent (blau) vorhergesagt wird. Rechts oben sind die Merkmalwerte (engl. Feature) angezeigt (z. B. »Merkmal 0« = 10). In der Mitte ist dargestellt, welche Merkmale für eine Anomalie und welche für ein Normalverhalten sprechen. Dabei ist zu erkennen, dass das »Merkmal 0«, welches der Wert eines bestimmten Ventils ist, am ausschlaggebendsten für die vorhergesagte Störung ist. Am zweitwichtigsten ist »Merkmal 4« für einen Hinweis auf eine Anomalie. Der Instandhalter bzw. die Instandhalterin der Anlage kann somit schrittweise die Liste der Merkmale durchgehen und die entsprechenden

Wartungsarbeiten durchführen (z. B. Ventil austauschen), um damit die vorhergesagte Störung zu vermeiden.

In Abbildung 2 ist das Ergebnis einer globalen Erklärung für ein KI-Verfahren dargestellt. Diese Darstellung vermittelt, auf welchen Merkmalen das Vorhersagemodell für die Anomalien beruht, die in einer bestimmten Anlage auftreten können. Dabei ist zu erkennen, dass das Auftreten einer Anomalie in den meisten Fällen mit dem »Merkmal 0«, wie auch bei der lokalen Erklärung, in Verbindung steht.

Nähere Informationen zum Thema Erklärbare KI und zur Toolbox können auf der Webseite: <https://www.iosb.fraunhofer.de/de/kompetenzen/bildauswertung/interaktive-analyse-diagnose/erklarbare-ki.html> gefunden werden.

KI-ENGINEERING FÜR DIE ÜBERWACHUNG VON DAMPF- UND GASTURBINEN

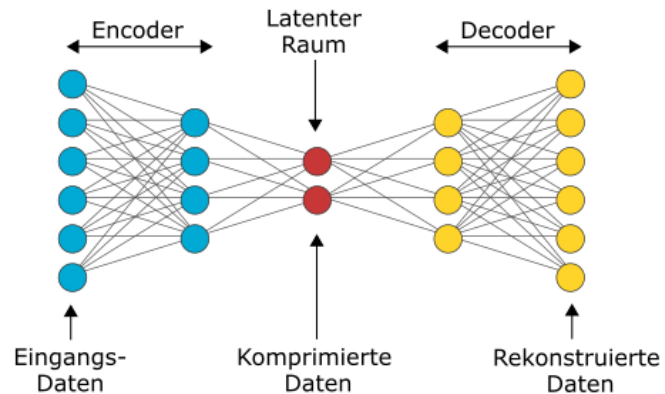


Abbildung 1: Beispiel eines Autoencoders.

Eine klassische Anwendung des Condition-Monitorings großer Dampf- und Gasturbinsätze basiert auf der hochaufgelösten Schwingungsanalyse der Turbinen im laufenden Betrieb. Die erfassten Schwingungsdaten werden dabei rechnergestützt durch Experten ausgewertet, und bei erkannten Abweichungen vom Normalbetrieb werden Handlungsoptionen abgeleitet, um Schäden am Prozess zu vermeiden. Im Zuge der Energiewende werden Großturbinen allerdings an Bedeutung verlieren und die Energieerzeugung wird sich auf kleinere, dezentralere Industrieturbinen verlagern. Dies wiederum bedeutet, dass Experten in Zukunft mehr Maschinen gleichzeitig betreuen und überwachen müssen.

Im Rahmen eines sogenannten QuickChecks (Projektlaufzeit eine Woche) wurde untersucht, inwiefern man eine auf KI bzw. maschinellem Lernen (ML) basierende Lösung entwickeln kann, um Experten bei der Überwachung der Turbinen zu unterstützen. Seitens des maschinellen Lernverfahrens lässt sich diese Fragestellung als eine Form der Anomaliedetektion auffassen. Bei der Anomaliedetektion wird ein ML-Modell auf einem zuvor ausgewählten Datensatz trainiert und dieses Modell wird laufend

mit aktuellen Messdaten abgeglichen. Wird in den Messdaten eine Abweichung zum Modell festgestellt, so wird davon ausgegangen, dass eine Anomalie im Prozess aufgetreten ist.

Zur Durchführung des QuickChecks wurden als maschinelles Lernverfahren ein Autoencoder (siehe Abbildung 1) und Messdaten im Normalbetrieb der Anlage zum Training genutzt. Im nächsten Schritt wurde der Autoencoder gegen einen Datensatz mit einer enthaltenen Störung, im Detail einer thermischen Verformung der Turbine, getestet. Mittels des trainierten Modells war es möglich, die im Prozess auftretende Störung korrekt zu erkennen und die Fehlerquelle anhand der relevanten Sensoren zu detektieren.

Die Ergebnisse haben gezeigt, dass mittels eines KI-gestützten Assistenzsystems eine vereinfachte Überwachung von Turbinen realisiert und Prozessexperten auf diese Weise unterstützt werden können. Zudem wurde im QuickCheck noch auf eine mögliche Systemarchitektur zur Umsetzung der KI-Methodik in räumlich entfernten Anlagen eingegangen und diese hinsichtlich Skalier- und Wartbarkeit der KI bewertet.



Dr.-Ing. Christian Kühnert

Mess-, Regelungs- und
Diagnosesysteme (MRD)
Fraunhofer IOSB Karlsruhe

Telefon +49 721 6091-511
christian.kuehnert@iosb.fraunhofer.de
www.iosb.fraunhofer.de/MRD

KI-ENGINEERING FÜR DIE SMARTE BAUSTOFFPRODUKTION

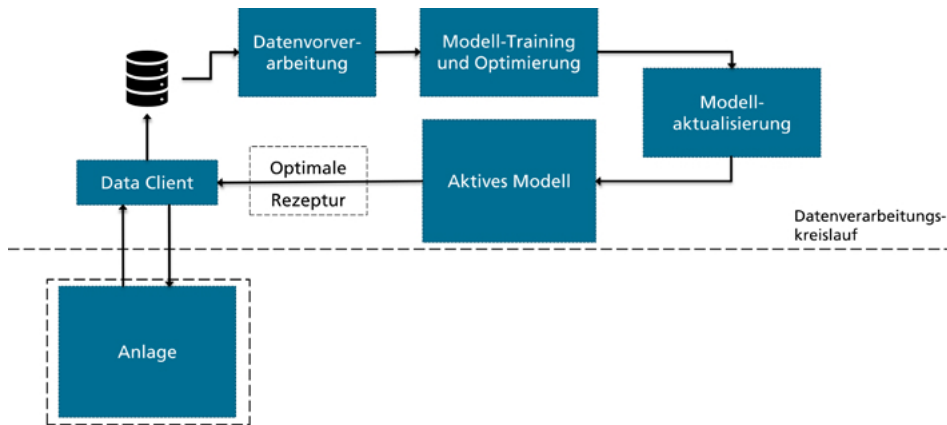


Abbildung 2: Bildentwurf für Datenkreislauf.

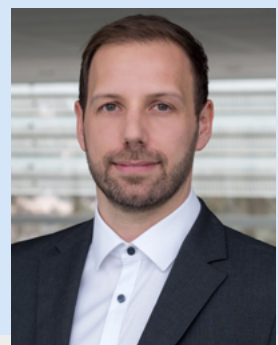
Da KI-Engineering nicht nur die Entwicklung eines KI-Modells zur Analyse von Daten aus einem bestimmten Anwendungsfall betrachtet, sondern die gesamte Spannweite von der Beschreibung einer Zielstellung über die Konzeptentwicklung bis zur fertig entwickelten Lösung abdeckt, wurde in einem weiteren QuickCheck ein grundlegendes Konzept für eine Fragestellung entwickelt.

Im Detail wurde untersucht, wie eine Systemarchitektur für die automatisierte und kontinuierliche Optimierung von Rezepturen von Baustoffen konzipiert sein muss. Die Baustoffe werden in Produktionsanlagen hergestellt, die während der Produktion mit einer Vielzahl von Sensoren die Daten aufzeichnen, aber auch kontrollierbare Aktuatoren, wie beispielsweise Heizelemente, beinhalten. Diese Kombination aus Sensorinformation und Steuerbarkeit der Anlage bildet die Grundlage für den QuickCheck.

Das entwickelte Konzept beschreibt einen Datenverarbeitungskreislauf, der sich an den Machine Learning Operations (MLOps)-Konzepten orientiert. In der Produktionsanlage werden kontinuierlich Produktionsdaten (Mengen, Temperaturen, Qualitätsindikatoren etc.) erfasst, diese werden

automatisiert verarbeitet und in einer Datenbank abgelegt. In definierten Zyklen werden basierend auf der wachsenden Datenbasis neue Modelle trainiert, die zur Optimierung der Rezepturen genutzt werden können. Vor dem Einsatz eines neuen Modells wird die Güte des Modells – anhand definierter Kriterien – mit dem schon eingesetzten Modell verglichen. Nur wenn eine Verbesserung der Modellgüte erkennbar ist, wird ein Austausch vorgenommen.

Im Zuge einer praktischen Umsetzung wird im nächsten Schritt eine funktionale Dekomposition des konzipierten Gesamtsystems vorgenommen und die einzelnen Subsysteme genauer beschrieben, hinsichtlich der an sie gestellten Anforderungen. Nachdem dies erfolgt ist, beginnt der Entwicklungszyklus des Systems. Wichtig ist, dass während des gesamten Ablaufs ein kontinuierlicher Austausch zwischen allen Beteiligten (Anlagenentwickler, Data Scientist, Softwareentwickler und Business Sponsor) erfolgen muss, um ein solches System realisieren zu können.



Christian Weißenbacher M.Sc.

Informationsmanagement
und Leitetchnik (ILT)
Fraunhofer IOSB Karlsruhe

Telefon +49 721 6091-531
christian.weissenbacher@iosb.fraunhofer.de
www.iosb.fraunhofer.de/ILT

KONTAKT

Themen



Dr. Masoud Roschani

Mess-, Regelungs- und
Diagnosesysteme (MRD)
Fraunhofer IOSB Karlsruhe

Telefon +49 721 6091-483
masoud.roschani@iosb.fraunhofer.de
www.iosb.fraunhofer.de/MRD



Jens Ziehn M.Sc.

Mess-, Regelungs- und
Diagnosesysteme (MRD)
Fraunhofer IOSB Karlsruhe

Telefon +49 721 6091-633
jens.ziehn@iosb.fraunhofer.de
www.iosb.fraunhofer.de/MRD

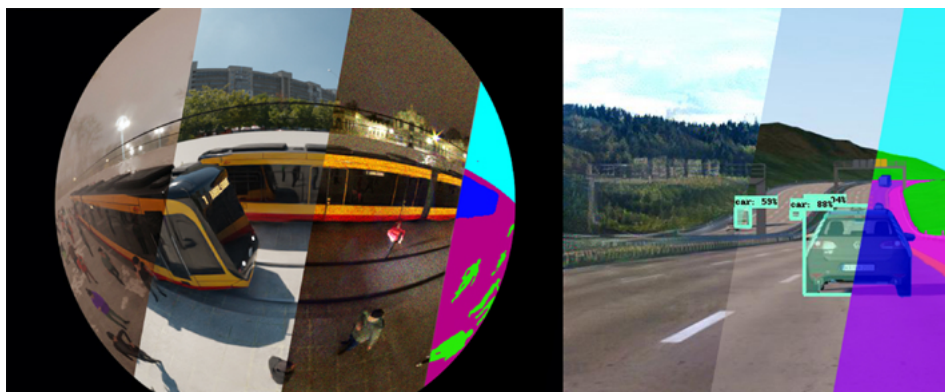
MASCHINELLES LERNEN UND MOBILITÄTSSYSTEME

Die Herausforderungen an Datenbedarfen und Interpretierbarkeit maschineller Lernverfahren und Künstlicher Intelligenz (KI), zu deren Lösung KI-Engineering einen Beitrag leisten soll, zeigen sich in besonderer Ausprägung im Bereich der automatisierten Mobilitätssysteme: Mit der Einführung hochautomatisierter Straßenfahrzeuge werden erstmals Systeme in Serie kommen, die in Alltagssituationen eigenständig hochsicherheitskritische Entscheidungen treffen – Entscheidungen, in denen maschinelles Lernen zunehmende Bedeutung gewinnt. Andere Anwendungen, wie automatisierte Schienenfahrzeuge und Flugtaxi, werden in absehbarer Zeit folgen.

Die zur sicheren Entwicklung und Erprobung notwendigen Datenumfänge werden dabei dramatisch steigen. Winner und Wachenfeld [1] führen etwa mit einer Beispielrechnung zu einer Größenordnung von 6,62 Mrd. Kilometern an Testfahrten, um nachzuweisen, dass ein Autobahnpilot sicherer fährt als der Mensch. Astronomische Größenordnungen wie diese – sie entspricht der Strecke von

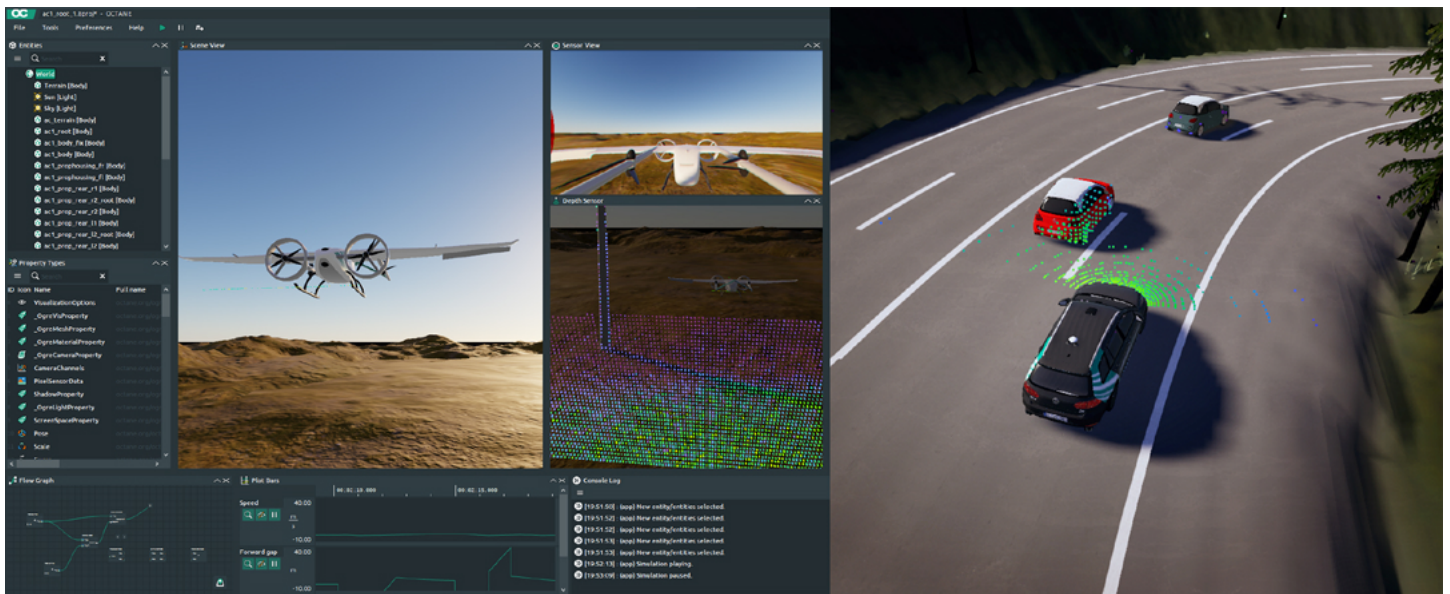
Karlsruhe bis zum Pluto – befördern die Relevanz von Simulationen, um Daten und Erprobungsszenarien bereitzustellen. Virtuelle Testfahrten sollen KI-Verfahren im öffentlichen Raum ohne reale Risiken an ihre Grenzen bringen. Sie sollen große Datenmengen ohne Datenschutzrisiken und Aufwände in der manuellen Nachbearbeitung bereitstellen und die Komplexität der »offenen Welt« durch Transparenz und Reproduzierbarkeit handhabbar machen. Dazu müssen Simulationen die Realität sowohl in ihrer Breite (das heißt: in ihrem Variantenreichtum) als auch in ihrer Tiefe (das heißt: in ihren wissenschaftlich erforschten Funktionsprinzipien) abbilden können. Denn alle unbekannten systematischen Verzerrungen in der Simulation können in diesem Einsatzbereich gravierende, reale Konsequenzen haben.

Aus diesem Grund wird am Fraunhofer IOSB – gemeinsam mit Partnern aus Fraunhofer-Gesellschaft, KIT, FZI und Industrie – die offene Simulationsplattform OCTANE (www.octane.org) entwickelt. Sie



Simulationsdaten aus der OCTANE-Plattform für KI-Anwendungen in der Mobilität: Detailgetreue Simulationen können helfen, KI-Verfahren schneller und zuverlässiger zu entwickeln und testen. Gleichzeitig können Simulationen selbst maschinelles Lernen nutzen, wie hier zur Synthetisierung von Texturen oder realitätsnaher Vegetation.

SIMULATION FÜR AUTOMATISIERTE



KI- und ML-Verfahren werden mittlerweile erfolgreich für diverse Anwendungen und Sensordaten eingesetzt. Diese müssen für unterschiedliche Fragestellungen in unterschiedlichen Detailgraden simulativ nachgebildet werden, um je nach Bedarf kurze explorative und interaktive Entwicklungszyklen zu unterstützen oder auch hochgenaue Detailberechnungen.

ermöglicht durch hohe Modularität und Flexibilität ein konsistentes Zusammenspiel interdisziplinärer Modelle: Von der Erzeugung virtueller Szenarien über die psychologische Verhaltenssimulation anderer Verkehrsteilnehmer bis hin zu Detailmodellen. Letztere umfassen beispielsweise Fahr- und Flugdynamik, Funksignal-Ausbreitung oder Sensorik wie Kameras und Laserscanner. Eine konsistente Austauschbarkeit von Modellen ermöglicht es – je nach Fragestellung – jeden Aspekt in dem Detailgrad zu simulieren, der tatsächlich erforderlich ist. Denn das eine, »perfekte« Simulationsmodell eines realen Prinzips gibt es nicht: Je nach Anforderungen und Rechenkapazitäten können leichtgewichtige, grobe Modelle sinnvoller sein, oder aber hochkomplexe und detaillierte Modelle.

Dabei greifen solche Modelle mitunter ihrerseits auf maschinelle Lernverfahren zurück – insbesondere dort, wo die Realität nur mit Mühe systematisch beschrieben werden kann. So wie KI-Verfahren lernen können, Bäume und Gesichter lediglich anhand vieler Beispiele zu erkennen (ohne Modelle biologischer Prinzipien zu erfordern), können sie auch lernen, realistische und vielfältige Bilder von Pflanzen und Personen zu *erzeugen*. Somit wird die Entwicklung von Simulationen erheblich darin beschleunigt, immer komplexere und umfangreichere Abbilder der realen Welt bereitzustellen. Auch hier zeigen sich Fragestellungen des KI-Engineerings, wenn es gilt, das Paradoxon zu lösen, wie Künstliche Intelligenz zuverlässig mit Verfahren erprobt werden kann, die selbst auf Künstlicher Intelligenz basieren.

Um die Welt der KI mit quantitativen, systematischen Entwicklungsprinzipien zu verbinden, legt OCTANE einige Prinzipien zugrunde: Die transparente und durchgängig physikalische Modellierung sowie den konsequenten Fokus auf wissenschaftlich belegte Modelle und Analysen. Das soll die Anwendbarkeit ebenso sicherstellen wie die Validierbarkeit der Ergebnisse – und somit Werkzeuge bieten, um vertrauenswürdige Abbilder der Realität für die ingenieurmäßige, interdisziplinäre KI-Entwicklung zu schaffen.

[1] H. Winner und W. Wachenfeld, „Die Freigabe des autonomen Fahrens“, in Autonomes Fahren - Technische, rechtliche und gesellschaftliche Aspekte, Springer, 2015, S. 458.

Fraunhofer-Institut für Optronik, Systemtechnik und Bildauswertung IOSB

info@iosb.fraunhofer.de

www.iosb.fraunhofer.de

STANDORT KARLSRUHE

Fraunhoferstraße 1
76131 Karlsruhe, GERMANY
Telefon +49 721 6091-0

STANDORT ETTLINGEN

Gutleuthausstraße 1
76275 Ettlingen, GERMANY
Telefon +49 7243 992-0

KONTAKTBÜRO BEIJING

Unit 0602G, Tower D1
DRC Liangmaqiao Office Building
19 Dongfangdonglu, Chaoyang District
100600 Beijing, PR China
Telefon +86 10 6590 0620
muh@fraunhofer.com.cn

Institutsteil für angewandte Systemtechnik Fraunhofer IOSB-AST

www.iosb-ast.fraunhofer.de

STANDORT ILMENAU

Am Vogelherd 90
98693 Ilmenau, GERMANY
Telefon +49 3677 461-0
info@iosb-ast.fraunhofer.de

AUSSENSTELLE GÖRLITZ

Wilhelmsplatz 11
02826 Görlitz, GERMANY
Telefon +49 3581 7925354
joerg.laessig@iosb-ast.fraunhofer.de

FORSCHUNGSGRUPPE SMART OCEAN TECHNOLOGIES ROSTOCK

Alter Hafen Süd 6
18069 Rostock, GERMANY
Telefon +49 3677 461-124
thomas.rauschenbach@iosb-ast.fraunhofer.de

Institutsteil für industrielle Automation Fraunhofer IOSB-INA

www.iosb-ina.fraunhofer.de

STANDORT LEMGO

Campusallee 1
32657 Lemgo
Telefon +49 5261 94290-22
juergen.jasperneite@iosb-ina.fraunhofer.de

