Fraunhofer

KI-Engineering in der Produktion

Whitepaper der Fraunhofer-Institute IOSB und IAIS

Thomas Usländer, Daniel Schulz (Hrsg.)

Fraunhofer Verlag

Fraunhofer-Gesellschaft

Thomas Usländer, Daniel Schulz (Hrsg.)

KI-Engineering in der Produktion

Whitepaper der Fraunhofer-Institute IOSB und IAIS

Fraunhofer Verlag

Kontakt:
Fraunhofer-Institut für Optronik, Systemtechnik und Bildauswertung IOSB
Fraunhoferstr. 1
76131 Karlsruhe
Telefon +49 721 6091-0
info@iosb.fraunhofer.de
www.iosb.fraunhofer.de

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.de abrufbar.

ISBN (Printausgabe): 978-3-8396-1943-8
DOI (kostenlose PDF-Version): https://doi.org/10.24406/publica-1685

Druck und Weiterverarbeitung: Elanders Waiblingen GmbH, Waiblingen

Für den Druck des Buches wurde chlor- und säurefreies Papier verwendet.

Nobelstraße 12
70569 Stuttgart
verlag@fraunhofer.de
www.verlag.fraunhofer.de

als rechtlich nicht selbständige Einheit der

Fraunhofer-Gesellschaft zur Förderung
der angewandten Forschung e.V.
Hansastraße 27 c
80686 München
www.fraunhofer.de

Herausgeber:
Dr.-Ing. Thomas Usländer
Fraunhofer-Institut für Optronik, Systemtechnik und Bildauswertung IOSB
Daniel Schulz
Fraunhofer- Institut für Intelligente Analyse- und Informationssysteme IAIS

Gefördert durch den Fraunhofer Cluster of Excellence »Cognitive Internet Technologies« CCIT.

Autorinnen und Autoren:

IOSB:
Christian Frey
Ann-Kathrin Goßmann
Dr. rer. nat. Constanze Hasterok
Philipp Hertweck
Dr. Christian Kühnert
Dr.-Ing. Julius Pfrommer
Dr.-Ing. Thomas Usländer

IAIS:
Dr. Gunar Ernis
Dr. Dirk Hecker
Dr. Maximilian Poretschkin
Daniel Schulz
Dr. Dennis Wegener
Dr. Tim Wirtz
Alexander Zimmermann

1 Vorwort

Die durch die Vision der Industrie 4.0 ausgelöste Dynamik hin zu vernetzten Maschinen und Anlagen sowohl innerhalb als auch zwischen produzierenden Unternehmen hat bislang das Thema der Interoperabilität zwischen Systemen in den Vordergrund geschoben. Und dies nicht nur in der physischen Welt der Maschinen, Anlagen und Produkte – den sogenannten Assets – sondern auch in der virtuellen Welt der digitalen Zwillinge dieser Assets. Durch die zunehmende Akzeptanz der Industrie 4.0 Standards kann der Aufwand der Vernetzung der Assets deutlich reduziert werden.

Der eigentliche wirtschaftliche Nutzen und das Werteversprechen der Industrie 4.0 zur Produktionsoptimierung, Produktverbesserung und neuen Geschäftsmodellen lässt sich aber nur erreichen, wenn die Daten der Assets optimal und automatisiert genutzt werden. Dazu eignen sich grundsätzlich Methoden der Künstlichen Intelligenz (KI).

KI-Anwendungen sind seit vielen Jahren ein wesentlicher Baustein in der angewandten Forschung bei Fraunhofer zur Lösung und Optimierung von Fragestellungen, die mit klassischen Methoden nur schwer greifbar und umsetzbar sind. Doch wie lässt sich die Erwartungshaltung der Ingenieure an eine prognostizierbare und quantifizierbare Leistungsfähigkeit von Systemlösungen zusammenbringen mit der zwar erstaunlichen, aber nicht sehr verlässlichen Lösungsfindung von datengetriebenen, KI-basierten Systemen?

Um diese Frage zu adressieren, bedarf es einer ingenieurmäßigen, systematischen Herangehensweise an die Nutzung von KI-Verfahren als Teil eines Systems Engineering-Prozesses von der Entwicklung bis hin zum operativen Betrieb. Wir nennen diese ganzheitliche Methodik KI-Engineering.

KI-Engineering ist ein Fokusthema des Fraunhofer Strategischen Forschungsfelds der Künstlichen Intelligenz. Die Fraunhofer Institute IOSB und IAIS arbeiten an der Definition und Umsetzung dieser entstehenden und aufstrebenden Disziplin sehr eng zusammen. Das vorliegende Whitepaper beleuchtet die wesentlichen Aspekte des KI-Engineering für die Domäne der industriellen Produktion.

Wir laden Sie ein zu einer spannenden Lektüre und freuen uns auf Ihre Rückmeldungen und Anregungen! Gerne arbeiten wir mit Ihnen gemeinsam an der Umsetzung des KI-Engineering in die Praxis!

Mit herzlichen Grüßen

Prof. Dr.-Ing. Jürgen Beyerer, Institutsleiter Fraunhofer IOSB
Prof. Dr. Stefan Wrobel, Institutsleiter Fraunhofer IAIS

Inhalt

2 Einführung

2.1 Motivation

Künstliche Intelligenz (KI) ist längst nicht mehr eine Vision der Zukunft, sondern hat sich bereits durch viele Anwendungen fest in unseren Alltag integriert. Spätestens seit der Veröffentlichung von OpenAI ChatGPT haben sich auch bislang eher zurückhaltende Unternehmen mit dieser Technologie befasst. Es ist daher nicht verwunderlich, dass in einer Studie von Deloitte 87 % der Unternehmen angeben, dass KI-Lösungen in den kommenden Jahren für den Gesamterfolg ihres Unternehmens eine wichtige bis sehr wichtige Bedeutung haben werden [1].

Doch obwohl die Bedeutung von KI für Unternehmen und Organisationen offensichtlich ist, kämpfen viele von ihnen immer noch mit der konkreten Umsetzung von KI-Technologien in ihre Arbeitsprozesse. So geben insgesamt 85 % der Unternehmen in einer Studie des Branchenverbandes Bitkom [2] an, dass sie entweder den Anschluss bei der KI-Umsetzung verpasst haben (42 %) oder sich zu den Nachzüglern zählen (43 %). Leider verbleiben auch viele KI-Projekte in Deutschland in einem Prototypenstatus und schaffen es nicht in einen erfolgreichen Produktivbetrieb.

In der industriellen Produktion sind die Optimierung und Steuerung von Prozessen, um Energie-, Material- oder Personalkosten zu sparen, oder die Vorhersage von Produktqualitäten, um Ausschuss zu vermeiden, zentrale KI-relevante Umsetzungsgebiete. Hier steht zumeist nicht wie bei der generativen KI die Kreativität der Lösung im Vordergrund, sondern eher die Verlässlichkeit und Vertrauenswürdigkeit der KI-basierten Lösung. Zukünftig verspricht aber auch in dieser Domäne der Einsatz der generativen KI spannende Lösungen (vgl. Ausblick).

Aber auch im industriellen Umfeld ist der Einzug von KI bisher noch gehemmt. Es sind zwar KI-Verfahren in den Unternehmen in das alltägliche Arbeitsumfeld eingezogen, jedoch konnten in der Gesamtbetrachtung laut einer Studie des VDI aus 2022 [3] die Erwartungen an ihren Einsatz bisher nicht erfüllt werden. Als Gründe hierfür werden fehlende Voraussetzungen, Datenmodelle und Systeme genannt. Um das unzweifelhafte Innovations- und Optimierungspotenzial von KI-Verfahren nutzen zu können und die genannten Hürden zu beseitigen, benötigt man eine dedizierte Methodik. Wir bezeichnen diese als KI-Engineering.

Das vorliegende Whitepaper stellt die technischen und organisatorischen Herausforderungen des KI-Engineering in der Produktion vor und gibt gleichzeitig praktische Handlungsempfehlungen für eine effiziente und erfolgreiche KI-Umsetzung. Hierzu beleuchten wir die vielfältigen Vorteile des KI-Engineerings, dessen Ziel es ist, KI- und ML-Methoden gemäß den typischen Anforderungen und Vorgehensweisen von Ingenieuren und Ingenieurinnen systematisch nutzbar zu machen – auch in sicherheitskritischen Anwendungen. Der Fokus der Veröffentlichung liegt daher in der industriellen Produktion als einem der bedeutenden Wirtschaftszweige in Deutschland.

2.2 KI-Engineering – Definition und Anwendungsdimensionen

Der Begriff KI-Engineering wird auf Englisch mit »AI Systems Engineering« übersetzt und ist wie folgt definiert [4]:

KI-Engineering adressiert die systematische Entwicklung und den Betrieb von KI-basierten Lösungen als Teil von Systemen, die komplexe Aufgaben erfüllen.

Die Ziele von KI-Engineering sind:

- Die Ermöglichung der Nutzung von KI im Rahmen der systematischen Herangehensweise von (Software-)Ingenieurdisziplinen.
- Die Entwicklung von Methoden, Werkzeugen und Prozessen, um die Entwicklung von

KI-Engineering ist eine bedeutende Methodik, um KI-Verfahren systematisch in technische Anwendungen zu integrieren. Die Erarbeitung entsprechender Normen und Standards ist essenziell für die Akzeptanz in den Unternehmen.«

Dipl.-Ing. Filiz Elmas, Leiterin Geschäftsfeldentwicklung Künstliche Intelligenz, DIN-Bereich Normung und Standardisierung

KI-Engineering Lösungen zu unterstützen. Dies beinhaltet eine formale Charakterisierung der Leistungsfähigkeit von KI-Lösungen zum Zeitpunkt der Entwicklung (im Gegensatz zu rein statistischen Betrachtungen der empirischen Leistung).
- Die Beschreibung einer neuen (Ingenieurs-) Disziplin, welche die Informatik sowie die datenbasierte Modellbildung und Optimierung mit dem Systems Engineering und den klassischen Ingenieurdisziplinen verbindet.

KI-Engineering stützt sich auf den ISO-Begriff eines KI-Systems als ein nach Ingenieurprinzipien entworfenes, »technisches System (engineered system), das Ergebnisse wie Inhalte, Vorhersagen, Empfehlungen oder Entscheidungen für eine bestimmte, vom Menschen definierte Zielsetzung erzeugt« [5].

Für den ganzheitlichen Ansatz des KI-Engineering wird davon ausgegangen, dass das Gesamtsystem hierarchisch in funktionale Komponenten (oder Subsysteme, je nach Komplexitätsgrad) zerlegbar ist. Diese können sowohl aus Hardware als auch aus Software bestehen. Ein oder mehrere Subsysteme können dabei KI-Systeme sein im Sinne der ISO-Definition. Wesentliche Bestandteile für die Funktionsfähigkeit des Gesamtsystems sind aber nicht nur KI-Subsysteme, sondern auch die Datensätze, welche für die Entwicklung der KI-Systeme benötigt werden. Zudem gibt es auch andere Subsysteme, die z. B. für die Sensordatenerfassung, Datenhaltung oder die Benutzerschnittstelle zuständig sind und nicht notwendigerweise KI-Verfahren enthalten. Diese gesamtheitliche Betrachtung ist ein zentraler Ansatz im KI-Engineering.

Die folgenden drei Dimensionen zeigen auf, wo der Unterschied liegt zwischen Anwendungen, die KI-Engineering benötigen und »allen anderen KI-Anwendungen«. Die Methoden, Werkzeuge und Prozesse, die im KI-Engineering entwickelt werden, können ebenfalls in Bereiche innerhalb dieser Dimensionen verortet werden. Eine einzelne Dimension kann bereits den Einsatz von KI-Engineering rechtfertigen.

Die Anwendung von KI-Verfahren in produktionstechnischen Systemen muss die Spezifika dieser Domäne berücksichtigen.

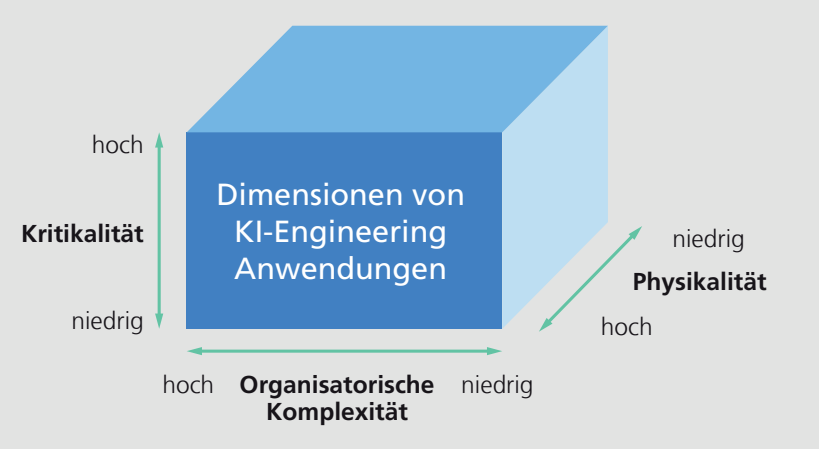


Bild 1: Dimensionen des KI-Engineering

In vielen Branchen (z. B. dem elektronischen Handel) entstehen durch den KI-Einsatz neue Geschäftsmodelle, so dass häufig neu entwickelt wird und Altsysteme oft nur angebunden werden müssen. In der industriellen Produktion ist es für die Systemarchitektur und die Infrastruktur ein wesentlicher Unterschied, ob völlig neue Fabriken mit einem neuen Maschinenpark entstehen (greenfield) oder ob KI-Verfahren in lange bestehende Produktionsanlagen mit bewährten, aber technologisch veralteten Maschinen integriert werden müssen (brownfield). Dies ist entscheidend für maschinennahe (edge computing) oder maschinenferne, cloud-basierte Datenverarbeitung (cloud computing), oder entsprechenden Mischformen mit all den Konsequenzen und Maßnahmen für die Datensicherheit und Verfügbarkeit.

Dimensionen des KI-Engineering [4]
- *Kritikalität*
  Bezieht sich auf die Auswirkungen eines nicht funktionierenden Systems auf Sicherheit (für Menschen und Systeme), geschäftskritische Funktionalität, Datenschutz und weitere Risiken.
- *Organisatorische Komplexität*
  Bezieht sich auf den Aufwand, 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 Unternehmen hinweg.
- *Physikalität*
  Diese Dimension 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. Diese Dimension ist ein Indikator für Kritikalität, jedoch sind nicht alle kritischen Anwendungen zwangsläufig in der physischen Welt verankert (man denke zum Beispiel an KI-basierte Angriffserkennung in der Cybersicherheit).

Aus systemtechnischer Sicht ist zudem die Frage der Interoperabilität auf allen Ebenen relevant, um flexible und zukunftssichere Systeme möglichst effizient entwickeln und betreiben zu können. Daher müssen die bestehenden und entstehenden Standards der Künstlichen Intelligenz [6], der Industrie 4.0 zur Datenmodellierung (z. B. die Industrie 4.0 Verwaltungsschale [7]) und der offenen Kommunikationssysteme wie z.B. IEC 62451 OPC UA berücksichtigt werden.

Die nachfolgenden beiden Tabellen verdeutlichen die wesentlichen weiteren technischen und organisatorischen Herausforderungen in der Entwicklung (linke Spalte) und im (langfristigen) Betrieb (rechte Spalte) von KI-Anwendungen).

Technische Herausforderungen	
in der Entwicklung	**im (langfristigen) Betrieb**
Leistungsprognose • garantierte Grenzen für die Leistung • frühzeitige Schätzung der Leistung in der Systementwurfsphase	Distribution Shift • Systeme sind im Laufe der Zeit nicht stabil • Modelle müssen angepasst werden, z.B. durch langsamen Abbau oder durch strukturelle Veränderungen • Um eine Verteilungsverschiebung zu erkennen, muss das eingesetzte Modell kontinuierlich überwacht werden
Datenverfügbarkeit • Daten aus dem Betrieb werden benötigt, bevor das System erstellt werden kann	Anlagenverfügbarkeit • Industrieanlagen werden oft über viele Jahrzehnte eingesetzt • Können wir in 10 Jahren einen Ersatz für genau diese GPU kaufen? • Sind Treiber noch auf einem unterstützten Betriebssystem verfügbar?
Heterogene Bereitstellung von Daten • feine Unterschiede zwischen den Betriebsbedingungen	Heterogene Bereitstellung von Daten • Qualität der Sensordaten • Heterogenität von ausgetauschten Sensoren
Integration von Domänenwissen • Kombination von Expertenwissen mit datengetriebenen Modellen	Nachhaltige Vermittlung von technischem Know-how • Sicherstellung, dass das organisatorische Wissen über das System über die Zeit erhalten bleibt

Organisatorische Herausforderungen	
in der Entwicklung	**im (langfristigen) Betrieb**
Lineare vs. iterative Entwicklungsmodelle • Systems Engineering bevorzugt die lineare Systementwicklung • KI-Entwicklung probiert verschiedene Lösungsansätze aus und kann sich nur schwer auf einen festen Zeitplan festlegen	Data Distribution Shift • Eingriffsmöglichkeit für die Anwender • Anlagenbetreiber reagieren negativ auf einen (gefühlten) Kontrollverlust • Erkenntnisse mit guter Erklärbarkeit und menschlichem Eingreifen • Alternative: keine menschlichen Bediener in der Kontrollschleife
Skalierung auf große, heterogene Teams	Benutzerfreundlichkeit • Differenzierung zwischen Gelegenheitsnutzer und Expertenanwender • Leistung abhängig von menschlichen Eingaben
Projektplanung und Risikomanagement • Dauer der KI-Entwicklung? • Wahrscheinlichkeit des Scheiterns? • Wird der richtige Lösungsansatz verfolgt? → Projektmanager, die keine KI-Experten sind, bewerten	Sichtbarkeit von Verbesserungen • Unterstützung des Managements bei KI-Lösungen durch Quantifizierung von Verbesserungen (im Laufe der Zeit)

3 Anwendungsfälle von KI-Engineering

Insgesamt bietet KI ein großes Potenzial für produzierende Unternehmen, um ihre Effizienz und Produktivität zu steigern und wettbewerbsfähig zu bleiben. KI umfasst jedoch nicht nur maschinelle Lernverfahren (ML), sondern beinhaltet auch klassische Verfahren wie die der Statistik und Optimierung. Dies ist ein wichtiger Aspekt, der häufig vernachlässigt wird. Nicht für jede Problemstellung sind ML-Verfahren am besten geeignet. Je nach Problemstellung können einfachere Verfahren gleichwertige oder sogar bessere Ergebnisse als ML liefern. Beispielsweise kann die Vorhersage der Ausfallwahrscheinlichkeit eines Bauteils zwar durch aufwändig trainierte neuronale Netze erfolgen, jedoch liefern in manchen Fällen deutlich weniger komplexe statistische Analysen vergangener Ereignisse ähnlich gute Ergebnisse. Ein anderes Beispiel für ein KI-Verfahren mit geringer Komplexität ist die Optimierung von Maschinenauslastungen durch simples Durchtesten möglicher Lösungen. Aber auch in diesen Fällen ist ein systematisches Vorgehen im Sinne des KI-Engineering sinnvoll.

Einer der ersten Schritte bei der Anwendung von KI besteht darin, vorhandene Daten zu nutzen. Häufig haben Unternehmen bereits Daten gesammelt, gewinnen aber aus diesen noch keine Erkenntnisse. Beispiele hierfür sind Sensordaten, die bereits abgespeichert oder Log-Nachrichten der Maschinen, welche in Datenbanken hinterlegt werden. Allein die Aufbereitung, Analyse und Visualisierung der vorhandenen Daten, z.B. in einem Dashboard kann hier schon einen Mehrwert bringen. So können beispielsweise ideale Wartungszyklen auf Basis von historischen Daten über Maschinenausfälle und -laufzeiten bestimmt werden. Bei Anwendungsfällen, bei denen noch nicht genügend Daten vorliegen und die Datengenerierung mit sehr hohem Aufwand verbunden ist bzw. zu erkennende Fehler nur selten vorkommen, kann die Datenbasis mit zusätzlichem Wissen aus anderen Datenquellen angereichert werden. Hierzu zählen z.B. Daten aus Simulationen, Wissensgraphen und das Domänenwissen von Fachexperten. Methodische Ansätze des Informed Machine Learning erlauben es, diese Daten in die Modellbildung zu integrieren [8].

Um ein passendes Lernverfahren auszuwählen ist es wichtig, die Größe der Datenmenge, auf der gelernt wird, und die Komplexität der Zusammenhänge zu berücksichtigen. Es gibt Lernverfahren, welche große Datenmengen benötigen (z.B. neuronale Netze) und solche, welche auch mit geringeren Mengen auskommen (z. B. Decision Trees und Support Vector Machines). Außerdem bestimmt die Art der Daten, welches Lernverfahren genutzt werden sollte. Während beispielsweise neuronale Netze für Bilddaten im Allgemeinen eine gute Performanz zeigen, sind Decision Trees häufig für tabellarische Daten, wie z. B. Daten mehrerer Sensoren (Spalten), zu einer Zeit (Zeile) besser geeignet.

Im Sinne des ganzheitlichen Ansatzes von KI-Engineering ist jedoch nicht allein das KI-Subsystem für das Verhalten und die Funktionalität des Gesamtsystems verantwortlich, sondern alle Komponenten, die mit dem KI-Subsystem (KI-Komponente) in Verbindung stehen. Relevant ist somit auch, wo die Daten entstehen und wie die Ergebnisse der KI-Komponente weiterverarbeitet werden. Hierbei spielen Autonomiestufen eine wichtige Rolle. Wir orientieren uns an den folgenden vier Stufen, wie sie im Kontext der Industrie 4.0 für industrielle Produktion definiert wurden [9].

- Stufe 1 sieht die KI-Komponente als Assistenzsystem für ausgewählte Funktionen. Der Mensch trifft jedoch noch alle Entscheidungen. Die Ausgabe der KI-Komponente wird nicht direkt mit den anderen Systemkomponenten gekoppelt, sondern wird vom Menschen verwendet, um Aktionen einzuleiten (human-in-the-loop-Ansatz). Der Mensch hat somit die volle Kontrolle und trägt die Verantwortung.
- In der Stufe 4 arbeitet das System autonom und adaptiv in vorgegebenen Systemgrenzen. Die Ausgabe der KI-Komponente wird nun direkt als Signal an andere Systemkomponenten weitergegeben und dort verarbeitet. Die Systemgrenzen sind insofern berücksichtigt, dass das System abschaltet, sofern die Grenzen überschritten werden. Der Mensch überwacht nur noch und greift in Notfallsituationen ein.

KI-Engineering integriert die Erfahrungen aus Jahrzehnten der Softwareentwicklung mit den Anforderungen der Operationalisierung von KI von Anfang an im Entwicklungsprozess. Damit wird sichergestellt, dass die wesentlichen Aspekte für den sicheren und verlässlichen Betrieb von KI-Anwendungen sowie deren Weiterentwicklung Bestandteil jeder Entscheidung sind. Es ermöglicht somit eine gleichermaßen zielgerichtete sowie effiziente Entwicklung von sicheren KI-Anwendungen.«

Martin Meßmer – Head of Data Pipeline – ZF Friedrichshafen AG

- Die Stufen 2 und 3 sind Abstufungen dazwischen, in denen die Autonomie der KI-Komponente von Stufe 1 zu Stufe 4 schrittweise gesteigert wird.

KI-Engineering ist nicht auf bestimmte Autonomiestufen ausgerichtet, sondern für alle Abstufungen einsetzbar. Die folgende Tabelle dient als praktischer Leitfaden für Anwendungsbeispiele des KI-Engineering und deren Kategorisierungen. Sie zeigt Beispiele für die Autonomiestufen 1 und 4 (wo passend) und ordnet den Problemstellungen die geeigneten und gängigsten KI-Verfahren zu.

Stufe	Anwendungsfall	Ziel	Eingangsgrößen	Ausgabegrößen	KI-Verfahren
Qualitätskontrolle					
1	Fehlerdetektion bei der Elektromotorenfertigung anhand von Vibrationssignalen (akkustisch, Körperschall) des Elektromotors im Betrieb [10]	Fehlerhafte Teile aussortieren	Vibrationssignale	Fehler ja/nein	Self-Organizing-Maps (SOM), Autoencoder (AE), Variational Autoencoder (VAE)
4	Automatische Endkontrolle von gerollten Metallbuchsen; Defektinspektion der Mantelfläche [11]	Fehler erkennen und einordnen	Kamerabild von Mantelfläche	Signal an Sortiermechanismus	Neuronales Netz (NN), Convolutional Neuronal Network (CNN
Fehlerzuordnung					
1	Verschiedene Fehler beim Laserschweißen (Schweißperle, Spritzer,…) anhand von Kamerabildern erkennen und unterscheiden [10]	Fehler erkennen und einordnen	Kamerabild der Schweißnaht	Fehlertyp	Neuronales Netz (NN), Convolutional Neuronal Network (CNN)
4	Automatisches Erkennen von Plastik im Bioabfall beim Recycling. Enthält der Bioabfall Plastik, wird er als Restmüll umdeklariert [12] [13]	Plastikanteil erkennen	Kamerabild des Mülls	Mülltyp (Bio-/Restmüll)	Neuronales Netz (NN), Convolutional Neuronal Network (CNN
Qualitätsprognose					
1	Elektrischen Widerstand einer Schweißnaht produziert durch Laserschweißen anhand der Prozessparameter vorhersagen [10]	Qualität der Schweißnaht bewerten	Prozessparameter	elektrischer Widerstand	Random Forest Regressor (RFR), Support Vector Regression (SVR), Ridge Regression, Lasso Regression
4	Vorhersage der Produktqualität von Kunststoffspritzteilen anhand von Prozessparametern und anschließende Sortierung [14]	Produktqualität sicherstellen	Druck und Temperaturverlauf	Signal an Sortiereinheit	Datenvorverarbeitung: Rohdaten verwenden oder Feature Extraction aus den Verlaufskurven (z.B. Polynom Fits, Min, Max, Mean, Ableitung, etc.); KI-Modell: Je nach Datenvorverarbeitung NN, RFR, SVR

Stufe	Anwendungsfall	Ziel	Eingangsgrößen	Ausgabegrößen	KI-Verfahren
Predictive Maintenance					
1	Frühzeitiges Erkennen von Lagerspiel oder Hydraulikproblemen bei Axialkolbenpumpen [15]	Wartungsbedarf frühzeitig erkennen	Geräuschsignale beim Betrieb der Pumpe	Warnung	Datenvorverarbeitung: Fourier Analyse; KI-Modell: Long Short Term Memory (LSTM), Recurrent Neural Network (RNN), AutoRegressive Integrated Moving Average (ARIMA)
4	Fehlerhaften Zustand eines Lasers vor Ausfall erkennen und Wartungsfirma automatisch informieren [16]	Wartung automatisiert initiieren	Betriebsparameter (z.B. Kühlwasserstand)	Mail an Wartungsfirma	KI-Modell: (LSTM), (RNN), (ARIMA)
Prozessüberwachung					
1	Unbekannte Anlagenzustände (Anomalien) bei der Holzwerkstoffplattenherstellung erkennen [11]	Anomalien frühzeitig beheben	Prozessparameter (Pressendruck, Leimmenge, Holzfeuchtigkeit, …)	Warnung bei Anomalien	SOM, Autoencoder, (VAE)
4	-> siehe Prozessregelung				
Prozessregelung					
1	Prozessoptimierung	Plastikanteil erkennen	Schnittparameter (Drehzahl, Vorschub,…)	optimale Parameter	RFR, SVR, Ridge Regression, Lasso Regression
4	Regelung von Gasturbinen um den Betrieb hinsichtlich Emissionen und Verschleiß zu optimieren und saisonale Feinjustierung zu reduzieren [18]	Prozessoptimierung	Schnittparameter (Drehzahl, Vorschub,…)	optimale Parameter	RFR, SVR, Ridge Regression, Lasso Regression
KI-gestützte Werkerassistenzsysteme					
1	Prognose der Bearbeitungszeiten im Blechzuschnitt [19]	Fertigungsplanung erleichtern	Auftragsdaten, 3D-Modell vom Bauteil	Bearbeitungszeit	Ensemble Modelle: eins für jedes Datenformat (Auftragsdaten -> Regressoren, 3D-Modell -> NN); Kombination der Modelle durch Boosting oder Stacking
4	Erkennung von Bauteilen und automatische Parametrisierung der Werkzeugmaschine	Durchlaufzeiten erhöhen und Fehlerquellen reduzieren	Kamerabild vom Bauteil	Maschinenparameter	Neuronales Netz (NN), Convolutional Neuronal Network (CNN

4 Qualitative Eigenschaften

KI hat ein großes Potenzial, um Produktionsprozesse zu automatisieren und zu verbessern. Während es relativ einfach ist, dieses Potenzial anhand von Prototypen zu demonstrieren, stellt die Integration in den operativen Betrieb mit all seinen Randbedingungen bisweilen eine Herausforderung dar. Dies betrifft insbesondere die nicht-funktionalen Anforderungen. Unter Bezug auf den ISO 25010 Standard zur Software-Produktqualität nennen wir diese auch qualitative Eigenschaften und meinen damit Produkteigenschaften wie z.B. die funktionale Angemessenheit, Effizienz, Kompatibilität, Verlässlichkeit, IT-Sicherheit, Vertrauenswürdigkeit, Wartbarkeit und die Transparenz. Etablierte Verfahren zur Softwareentwicklung und -Qualitätssicherung greifen aber für KI-Systeme, die auf Maschinellem Lernen beruhen oftmals zu kurz, so dass entsprechende Erweiterungen für die Qualitätssicherung von KI-Systemen notwendig werden [20]. In diesem Zusammenhang hat sich der Begriff der Vertrauenswürdigkeit oder auch Akzeptierbarkeit von KI-Systemen herausgebildet.

Gemäß der von der EU-Kommission eingesetzten hochrangigen Expertengruppe für Künstliche Intelligenz (HEG-KI) ist für die Vertrauenswürdigkeit insbesondere die technische Robustheit und Sicherheit zu nennen [21]. Diese umfasst

- die Widerstandsfähigkeit gegen Angriffe und Sicherheitsverletzungen,
- einen Auffangplan z.B. den Umstieg von einem statistischen Verfahren auf ein regelbasiertes Verfahren oder Übergabe an einen menschlichen Bediener,
- die allgemeine Sicherheit gemäß dem von einem KI-System ausgehenden Risiko,
- die Präzision, d.h. die Fähigkeit eines KI-Systems, Sachverhalte richtig zu beurteilen, sowie
- die Zuverlässigkeit und Reproduzierbarkeit des Verhaltens gemäß der Spezifikation auch bei unterschiedlichen Eingaben und Situationen.

Darüber hinaus muss das KI-Engineering in der Produktion von vornherein auch auf den langfristigen Betrieb von KI-basierten Lösungen in der Industrie ausgerichtet sein. Dies erfordert vor allem die Akzeptierbarkeit und Erklärbarkeit KI-basierter Lösungen als Grundvoraussetzung für deren Akzeptanz durch die Anlagenbetreiber.

Die Anforderungen der HEG-KI sind in der Zwischenzeit durch die geplante KI-Verordnung [22] weiter konkretisiert worden. Diese verfolgt einen risikobasierten Ansatz und stuft Systeme, welche Sicherheitsfunktionen haben, als Hochrisikosysteme ein, die vor dem Inverkehrbringen einer Konformitätsbewertung unterzogen werden müssen. Neben der horizontalen KI-Verordnung unterliegt der Einsatz von KI in der Produktion weiteren regulatorischen Anforderungen. Zu nennen sind hier

- Gesetze, welche den Zugang bzw. die Verwendung von Daten sowie die entsprechenden Infrastrukturen adressieren (z.B. der Data Act, der Data Governance Act, Digital Services Act, Digital Markets Act),
- Gesetze, die Datenschutz und Cybersicherheit adressieren (z.B. der geplante Cyber Resilience Act), oder
- Vorgaben zur Produkthaftung und Sicherheit (etwa die Produkthaftungsrichtlinie und die Maschinenverordnung, welche die Maschinenrichtlinie ablöst).

Dabei ergänzen sich die Gesetze gegenseitig: Die Maschinenrichtlinie adressiert beispielsweise primär die Sicherheitseigenschaften eines Gesamtsystems (welches KI-Verfahren in Subsystemen enthalten kann), während die KI-Verordnung unmittelbar die KI-Komponenten im Blick hat, die die sicherheitskritischen Funktionen einer Maschine steuern.

Eine wichtige Rolle zur Operationalisierung der regulativen Anforderungen nimmt die Transparenz von KI-Systemen ein: Hier sind insbesondere die Rückverfolgbarkeit und Erklärbarkeit von Entscheidungen von elementarer Bedeutung. Gemäß [21] bezieht sich die Erklärbarkeit auf die Möglichkeit, sowohl

die technischen Prozesse eines KI-Systems als auch die damit verbundenen menschlichen Entscheidungen (z.B. Anwendungsbereiche eines KI-Systems) zu erklären. Technische Erklärbarkeit setzt voraus, dass die von einem KI-System getroffenen Entscheidungen vom Menschen verstanden und rückverfolgt werden können. Darüber hinaus müssen möglicherweise Kompromisse zwischen einer verbesserten Erklärbarkeit eines Systems (was die Präzision beeinträchtigen kann) und mehr Präzision (auf Kosten der Erklärbarkeit) eingegangen werden.

Für die Beurteilung der Erfüllung dieser Anforderungen ist es wesentlich, den Systemrahmen zu betrachten. Wie oben beschrieben, sind die KI-Verfahren oftmals in eigenständigen Subsystemen eingebaut, die auch in der Konzeption und Entwicklung gesondert behandelt werden. In einem KI-Engineering Vorgehensmodell, wie z.B. PAISE® (vgl. Kapitel 6.5), werden KI-Subsysteme deshalb typischerweise agil entwickelt und über Checkpoints mit anderen Subsystemen iterativ integriert und versioniert. Zudem spielen die KI-Subsysteme als Hilfssysteme (enabling systems) in der Entwicklungsumgebung eine wesentliche Rolle, zusammen mit denen für ML-Verfahren unabdingbar notwendigen (Trainings-) Datensätze.

KI-Engineering umfasst im klassischen Sinne des Systems Engineering das Design des Gesamtsystems, da letztlich dieses die entscheidenden Qualitätskriterien erfüllen muss. Auch ist es das Gesamtsystem, das in kritischen Umgebungen ggf. von einer Prüfinstanz abgenommen und geprüft wird, nicht nur die KI-Verfahren selbst. Als Grundlage für solche Prüfungen und perspektivisch auch KI-Zertifizierungen müssen entsprechende Prüfkriterien und Standards ausgearbeitet werden, einen konkreten Ansatz stellt der KI-Prüfkatalog [23] dar. Neben regulativen Anforderungen und der Erfüllung bestehender Sicherheitsstandards ist die Frage von KI-Qualität und der Nachweisführung der hinreichenden Mitigation KI-spezifischer Risiken wesentlich für die Etablierung neuer Geschäftsmodelle im Produktionsumfeld, etwa für KI-Gütesiegel, als Grundlage für KI-Versicherungen und Risikotransfers oder im Rahmen von Due Diligence Prüfungen.

5 Technische und organisatorische Schulden und Lösungsansätze

Software-Anwendungen, welche KI-Verfahren beinhalten, sind einem ständigen Wandel unterworfen. Um diese Anwendungen langfristig nutzen zu können, müssen sie kontinuierlich angepasst, aktualisiert und weiterentwickelt werden. Dies gilt auch für das Wissen der Mitarbeitenden in den Unternehmen. Daher gilt es, durch den Einsatz von KI-Engineering sowohl technische als auch organisatorische Schulden möglichst gering zu halten.

- Technische Schulden: Diese beziehen sich konkret auf das zu entwickelnde KI-System. Sie beinhalten die negativen Eigenschaften des Systems, die zum einen durch schlechte und lückenhaft durchgeführte Programmierarbeit und zum anderen durch im früheren Projektzeitraum getroffene Entscheidungen entstehen. Sie wirken sich üblicherweise negativ auf Qualitätseigenschaften wie Performance, Stabilität oder Wartbarkeit der Software, aus.
- Organisatorische Schulden: Diese ergeben sich, wenn Unternehmen notwendige Änderungen hinsichtlich ihrer Struktur, ihrer Prozesse oder ihrer Kultur nicht durchführen (können). Hinsichtlich KI-Engineering gilt dies vor allem, wenn es Widerstände gegen Veränderungen im Unternehmen gibt oder auch das benötigte Fachpersonal bzw. Fachwissen für KI-Projekte nicht vorhanden ist.

KI-Engineering strebt daher an, dass direkt zu Beginn der Systementwicklung sowohl auf eine strukturierte Arbeitsweise als auch auf geeignete Werkzeuge geachtet wird. Es gilt zu beachten, dass sowohl technische als auch organisatorische Schulden auf den späteren Erfolg bzw. Misserfolg von KI-Systemen ähnlichen Einfluss haben und im Vergleich zu klassischer Software die Integration einer KI-Komponente in das System den Softwareentwicklungsprozess noch komplexer werden lässt. Denn neben den Problemen der Softwareentwicklung steht ein KI-System zusätzlich vor der Herausforderung, dass nicht alle Eigenschaften der KI im Vorfeld bekannt sind und erst iterativ während des Projektverlaufs erarbeitet werden müssen. Zudem besteht bei KI-Systemen der ständige Bedarf, Datensätze und KI-Modelle in regelmäßigen Abständen zu aktualisieren. Denn sobald sich der grundlegende Datensatz ändert, müssen die Modelle neu gelernt und wieder in Betrieb genommen werden. Für den Erfolg eines KI-Projektes ist daher eine querschnittliche Datenstrategie zur Bereitstellung von Daten gleicher Qualität in allen Entwicklungsstadien äußerst wichtig. Zudem sollten mögliche rechtliche Fragestellungen im Vorfeld geklärt und berücksichtigt werden. Des Weiteren hat sich gezeigt, dass während der Entwicklung von KI-Systemen oftmals unter der Oberfläche schlummernde technische/organisatorische Schulden des Unternehmens offensichtlich werden, die viel umfangreicher sind als zu Projektbeginn erwartet. Klassische Beispiele sind der Wechsel von auftragsorientierter (batch) hin zu ereignisorientierter Datenverarbeitung oder der zukünftige Einsatz von Microservices anstelle von monolithischer Software. Zuletzt darf nicht vergessen werden, dass sich die Methoden- und Tool-Landschaft im Bereich KI sehr schnell entwickelt. Es lassen sich daher nur bedingt Entscheidungen für einen langfristigen und nachhaltigen Einsatz solcher Tools treffen. Daher kommt es hier umso mehr auf die Definition und den Einsatz übergreifender Prozesse und einheitlicher KI-Artefakte an.

Um eine kontinuierlich hohe Qualität des KI-Systems sicherstellen zu können, fordert KI-Engineering, dass bei allen Beteiligten ein Bewusstsein für technische und organisatorische Schulden geschaffen wird. Dies ist Grundvoraussetzung, um daraus resultierende Risiken zu erkennen, zu vermeiden und beseitigen zu können. Durch geeignete Maßnahmen sollte bereits zu Beginn einem Anhäufen von Schulden entgegengewirkt werden. Sind diese bereits entstanden, bzw. nicht zu vermeiden, so gilt es diese sichtbar zu machen, um allen Beteiligten die dadurch entstehenden Risiken transparent zu machen. Dies erlaubt eine bewusste Entscheidung darüber, ob die Schulden weitergetragen werden können, oder ob entsprechende Gegenmaßnahmen zur Risikoreduktion eingeleitet

werden müssen. Im Folgenden sind einige technische und organisatorische Schulden, speziell für KI-Systeme aufgeführt und es werden Lösungsvorschläge zu deren Beseitigung gegeben.

Technische Schulden in KI-Systemen	Mögliche Lösung
Unzureichende Testabdeckung; vor allem fehlende End-to-End Tests des KI-Systems	Durchgängige, automatisierte Tests, einzelner Softwaremodule sowie Testen des Gesamtsystems (unter Berücksichtigung der Testpyramide)
Ignorieren von Best Practices in der Softwareentwicklung	Etablierung von Programmierrichtlinien und Handlungsempfehlungen zur Strukturierung und Formatierung des Quellcodes. Entwicklung von Richtlinien speziell für KI-Systeme
KI-Lösung ist nicht ausreichend skalierbar	Entwicklung einer adäquaten Systemarchitektur der KI-Lösung. In Produktionssystemen vor allem auch Berücksichtigung des Edge-Cloud-Kontinuums
ML-Modell ist mit weiteren Softwaremodulen stark gekoppelt	Korrekt durchgeführtes Bereitstellen von ML-Modellen mit Entkopplung von weiteren Softwaremodulen
ML-Modell wird hinsichtlich Performanz und Zuverlässigkeit nicht überwacht	Ständiges Monitoring der ML-Modelle und z.B. Überprüfung hinsichtlich zeitlicher Änderungen der Datenverteilungen
ML-Modell wird als reines Black-Box-Modell genutzt	Nutzung transparenter und erklärbarer ML-Modelle und hierdurch Sicherstellung, dass Ergebnisse der Modelle keinen Bias aufweisen.
Datengrundlage für das ML-Modell wird nicht hinterfragt	Ständig die Anforderungen an die technologische Souveränität sowie die Datensouveränität für das zu entwickelnde KI-System beachten
Unzureichende Integration von Domänenwissen Unzureichende Kombination von Expertenwissen mit datengetriebenen Modellen	Nachhaltige Vermittlung von technischem Know-how Sicherstellung, dass das organisatorische Wissen über das System über die Zeit erhalten bleibt.

Organisatorische Schulden in KI-Systemen	Mögliche Lösung
Keine Klarheit, welches Teammitglied welche Rolle in der Entwicklung der KI-Lösung besitzt	Definitionen klarer Rollen und Zuständigkeiten. Hierdurch Abbau von Fehlkommunikation und Doppelarbeit
Einzelne Teammitglieder haben keinen Überblick über die Gesamtlösung	Regelmäßige Status-Updates und Projektmeetings. Alle Team-Mitglieder sollten jederzeit über Fortschritte und Probleme des KI-Systems informiert sein
Speziell mit Blick auf die KI baut das Team auf veralteter Technologie auf	Bereitstellung von Schulungs- und Ausbildungsmaßnahmen für die Teammitglieder. Sicherstellung, dass diese über die für ihre Rolle zugewiesenen Fähigkeiten für das KI-Systems verfügen
Unternehmen beginnt erstmalig mit der Entwicklung eines KI-basierten Systems	Ständiges Hinterfragen und Verbessern bestehender Prozesse im Unternehmen mit dem Ziel, die Gesamteffizienz des KI-Projektes zu verbessern

6 Vorgehensmodelle für KI-Engineering

MLOps ist die konsequente Weiterentwicklung von DevOps Prinzipien auf die Herausforderungen bei der Entwicklung und dem Einsatz von KI-Anwendungen. Dass das richtig ist, zeigt sich bereits dadurch, dass diese junge Disziplin in kurzer Zeit bereits sehr viele Anhänger hinter sich versammelt hat und mittlerweile der defacto Standard für effiziente Entwicklung von sicheren KI-Anwendungen ist.«

Sebastian Bader – Product Owner at SAP

Der Schlüssel, um die oben genannten organisatorischen Herausforderungen zu adressieren und Lösungen für die technischen Herausforderungen systematisch zu erarbeiten, ist ein Vorgehensmodell für KI-Engineering als Gesamtkonzept. Dies bedeutet, dass ein systematischer Weg aufgezeigt wird von der Zieldefinition eines Gesamtsystems bis hin zur ersten Installation und deren Weiterentwicklung und Pflege im Betrieb.

Dies findet statt im Spannungsfeld zwischen den Methoden des maschinellen Lernens, der Software-Entwicklung und dem Systems Engineering sowie dem Anwendungsfeld der industriellen Produktion. Zu allen Bereichen gibt es bereits etablierte Vorgehensmodelle und Entwicklungsmethoden, die allerdings nur Teilaspekte im Sinne des KI-Engineering abbilden. Eine Auswahl an Vorgehensmodellen wird in den folgenden Unterkapiteln kurz erläutert, bevor dann mit PAISE® das Vorgehensmodell für KI-Engineering vorgestellt wird.

6.1 DevOps

DevOps wird als Überbegriff verwendet, um Methoden und Werkzeuge der Entwicklung (development) und des IT-Betriebs (operations) von Software-Lösungen in einen kontinuierlichen Prozess zu integrieren. Wesentlich ist hierbei auch eine Kultur der Zusammenarbeit zwischen den Menschen in beiden Bereichen, die in vielen Unternehmen organisatorisch oftmals getrennt sind. Bei DevOps können die Software-Entwickler auch für den IT-Betrieb zuständig sein.

Das zugrunde liegende iterative Vorgehen leitet sich aus Prinzipien agiler Entwicklungsmethoden ab:

- Die handelnden Individuen und deren Interaktion werden höher bewertet als vordefinierte Prozesse und Werkzeuge.
- Funktionierende Software ist wichtiger als Dokumentation.
- Zusammenarbeit mit dem Kunden ist wichtiger als Vertragsverhandlungen über Produkteigenschaften.
- Reaktionsfähigkeit auf Änderungen in den Anforderungen sind wichtiger als Planerfüllung.

DevOps zielt auf Software-Lösungen unabhängig von der Methodik ihrer Komponenten ab. Einerseits passen gerade diese Prämissen sehr gut für die Entwicklung von KI-Subsystemen, andererseits widersprechen sie oft den Bedürfnissen einer spezifikations- und dokumentationsgetriebenen Entwicklung kritischer Systeme, welche im KI-Engineering im Vordergrund steht. Deshalb werden hierfür Kompromisslösungen gesucht. Gerade im Bereich der Software-Intensiven Systems of Systems (SISOS) kommt es seit den 2010er Jahren zu einer Rückbesinnung. Standards wie ISO 12207 und ISO/IEC 15288 ebnen dabei den Weg für die Entwicklung immer komplexerer Systeme, so wie es beim KI-Engineering das Ziel ist.

6.2 CRISP-DM und ASUM-DM

Das wohl bekannteste Vorgehensmodell im Bereich des Data Mining ist der Cross-Industry Standard Process for Data Mining (CRISP-DM) [24], welcher bereits 1996 definiert wurde. Data Mining wendet systematisch statistische Methoden an, um in Datenbeständen versteckte Muster und Strukturen aufzudecken. Mit CRISP-DM werden in sechs Phasen iterativ so lange neue Modelle (Prototypen) entwickelt, bis diese zufriedenstellend sind und in die Auslieferung gebracht werden können. Dieser Ansatz eignet sich auch für die Entwicklung von ML-Komponenten.

IBM veröffentlichte 2015 mit ASUM-DM (Analytics Solutions Unified Method for Data Mining/Predictive Analytics) [25] eine Überarbeitung und Erweiterung von CRISP-DM. Es wurde in ASUM-DM insbesondere eine Phase »Operate & Deploy« mit aufgenommen. Weiterhin wird der iterative Ansatz von CRISP-DM erweitert und ein lineares Projektmanagement mit iterativen Entwicklungsschleifen

verbunden. Interessant für das KI-Engineering ist, dass ASUM-DM den Fokus auf Optimierungsschritte nach der (Erst-) Installation (deployment) legt, also ausdrücklich den operativen Betrieb wie bei DevOps miteinschließt und unterstützt.

6.3 MLOps

MLOps vereint den DevOps-Ansatz, der die Zusammenarbeit zwischen Softwareentwicklung und IT-Betrieb regelt und unterstützt, mit den Komponenten des CRISP-DM [24]. Der Fokus liegt dabei auf den ML-spezifischen Anteilen, die sich von gängigem Software-Engineering aufgrund der Beschaffenheit der Methodik und/oder der Artefakte unterscheiden. MLOps ist wie DevOps mehr als eine Sammlung technischer Methoden. Es schließt ausdrücklich die Organisation, deren Prozesse und Personal mit ein.

MLOps deckt den gesamten Lebenszyklus einer KI-basierten Lösung ab. Es startet also mit der ersten Anforderungsanalyse, über die iterative Entwicklung und den dauerhaften Betrieb bis zu einer etwaigen Ablösung/Deaktivierung. Der MLOps-Zyklus beschreibt den iterativen Entwicklungsvorgang einer KI-basierten Lösung, von der Anforderungsanalyse bis zum Betrieb. Einzelne Bestandteile des MLOps-Zyklus kommen daher nicht einmalig, sondern wiederkehrend vor, allerdings in ggf. unterschiedlichen Laufzeitumgebungen. Dies gewährleistet eine passgenaue Unterstützung der Aktivitäten durch geeignete technische Werkzeuge und organisatorische Strukturen.

Wie bei CRISP-DM und ASUM-DM steht bei MLOps vor allem das Zusammenspiel aller beteiligten Stakeholder aus der Fachseite, der Entwicklung und des Betriebs im Vordergrund. In Bezug auf das Gesamtsystem, das für das KI-Engineering ausschlaggebend ist, liegt der Fokus allerdings auf der oder den KI/ML-Komponente(n).

6.4 V-Modell

Das V-Modell ist ein gängiges Vorgehensmodell im Systems Engineering, d.h. es nimmt von vornherein das Gesamtsystem in den Blick, basierend auf einem hierarchischen Systemmodell. Es wurde primär als lineares Vorgehensmodell entwickelt, lässt jedoch auch iterative Zyklen zu. Das V-Modell besitzt eine starke Dokumentenorientierung und eine V-förmige Gegenüberstellung der Entwicklungs- und Testphasen, jeweils bezogen auf Subsysteme und deren Komponenten. Ergebnisse einer Phase sind bindend für die nächste Phase. Agile Prozesse wurden mit dem V-Modell XT eingeführt, während in der neuesten Fassung des V-Modells [26] auch die Belange von cyber-physischen Systemen berücksichtigt werden.

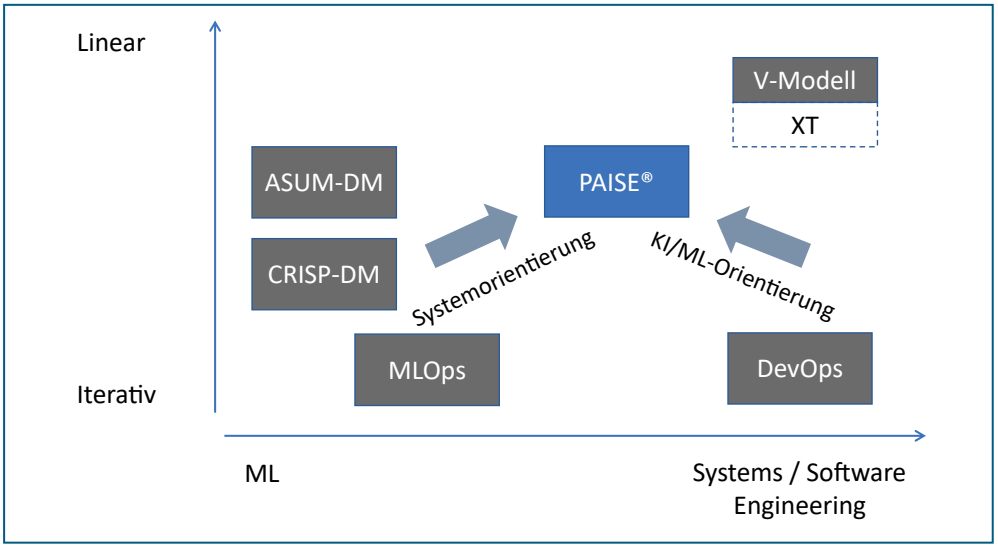


Bild 2: Einordnung des KI-Engineering Vorgehensmodells PAISE®

MLOps fördert die Zusammenarbeit zwischen unseren Data Scientists, Entwicklern, Produktmanagern und Schadensabwicklungs-Experten über verschiedene Abteilungen hinweg. Es stellt eine Struktur für die effektive teamübergreifende Zusammenarbeit und den Austausch von Modellen, Daten und Code bereit. Dadurch können unsere Teams schneller und effizienter arbeiten. Prototypen können bei uns so schnell in die konkrete Anwendung gebracht werden.«

Dr. Sebastian Schoenen – Head of Research and Development at ControlExpert GmbH, Teil der Allianz SE

Als Maschinenbauer sehen wir durch das KI-Engineering einen Mehrwert für unsere Kunden und einen technologischen Vorteil durch die Integration von KI-Verfahren in die Steuerungs- und Datentechnik der automatisierungstechnischen Anlagen. Das Vorgehensmodell PAISE® hilft uns, zielgerichtet die besten KI-Ansätze zu finden und zeitnah in eine praktische Engineering-Lösung zu überführen.«

Horst Fritz, Geschäftsführer Fritz Automation GmbH, Forbach

6.5 PAISE® (Process Model for AI Systems Engineering)

Ein Vorgehensmodell für KI-Engineering muss sowohl die Systemorientierung als auch die KI/ML-Orientierung beinhalten. Da die oben beschriebenen etablierten Modelle immer nur Teilaspekte abdecken, benötigt man für das KI-Engineering einen integrierenden, übergreifenden Ansatz, der das Gesamtsystem in der Entwicklung und im Betrieb im Blick hat und KI/ML-Verfahren in Subsystemen explizit unterstützt.

Mit diesem Ziel wurde im Kompetenzzentrum für KI-Engineering CC-KING (vgl. Kapitel 8.2) das Process Model for AI Systems Engineering PAISE® konzipiert. PAISE® betrachtet die Entwicklung eines Gesamtsystems, das in Subsysteme zerlegbar ist, die KI-Systeme sein können [27]. Das folgende Schaubild zeigt die sechs linear angeordneten Phasen des PAISE® Vorgehensmodells, wobei in der Phase »Entwicklungszyklus« die Entwicklungs- und Verfeinerungsschritte zyklisch durchlaufen werden. Gestützt und reguliert wird dies durch Checkpoints, in denen eine (Teil-) Integration und Bewertung der Komponenten hinsichtlich der Anforderungen stattfindet. Durch diesen zyklischen Prozess wird ein Wechsel zwischen explorativem Vorgehen auf der einen Seite und zielgerichtetem Vorgehen auf der anderen Seite möglich.

Dieser Ansatz adressiert eine wesentliche Herausforderung beim Einsatz von ML-Verfahren: Es bestehen inhärente Risiken beim ML-Entwicklungsprozess, welche sich z.B. aus der Abhängigkeit von der Datenqualität, die zu einer Einschränkung der Funktionalitäten führen kann, oder der Nichtvorhersagbarkeit der Performanz von KI-basierten Systemen, ergeben. Somit wird in PAISE® ein risikobasiertes Vorgehen verfolgt. Das bedeutet entsprechend, dass die folgenden Aspekte in PAISE® wichtig sind:

- Festlegung von Zielen (und für ML-basierte Verfahren wichtige Zielmetriken), Identifikation von Alternativen (KI-basiert vs. klassisch) und Beschreibung von Rahmenbedingungen.
- Erkennen, abschätzen und reduzieren von Risiken, z. B. durch Analysen, Simulationen oder Prototyping.
- Realisierung und Überprüfung des Zwischenprodukts.
- Planung des nächsten Zyklus der Projektfortsetzung.

Ein Hauptunterschied von PAISE® zu CRISP-DM ist, dass bei CRISP-DM die Entwicklung von datengetriebenen Modellen im Vordergrund steht. Die Existenz der Daten wird vorausgesetzt. PAISE® betrachtet Datensätze als gleichwertige Komponenten zu KI-Komponenten und berücksichtigt so deren Abhängigkeiten bei der Entwicklung. Ebenso wird die gleichwertige Entwicklung von datengenerierenden KI-Komponenten betrachtet. Somit werden Datenverfügbarkeit und Qualität innerhalb eines Systems bestehend aus Subsystemen mitentwickelt.

PAISE® sieht Tests unter Nutzung aller bestehenden Sub- und Hilfssysteme (auch in Form von Simulationen) vor und geht damit noch weiter als die Deployment-, Test- und Betriebsstrategie aus MLOps. Zudem berücksichtigt PAISE® auch stark einschränkende Randbedingungen, wie beschränkte Rechen- und Speicherressourcen, dynamische Interaktionen sowie Abhängigkeiten aller Subsysteme. Grundlage bietet hierfür die Komponentenspezifikation, welche funktionale und nichtfunktionale (qualitative) Anforderungen (vgl. Kapitel 4) klar benennt und modularisiert.

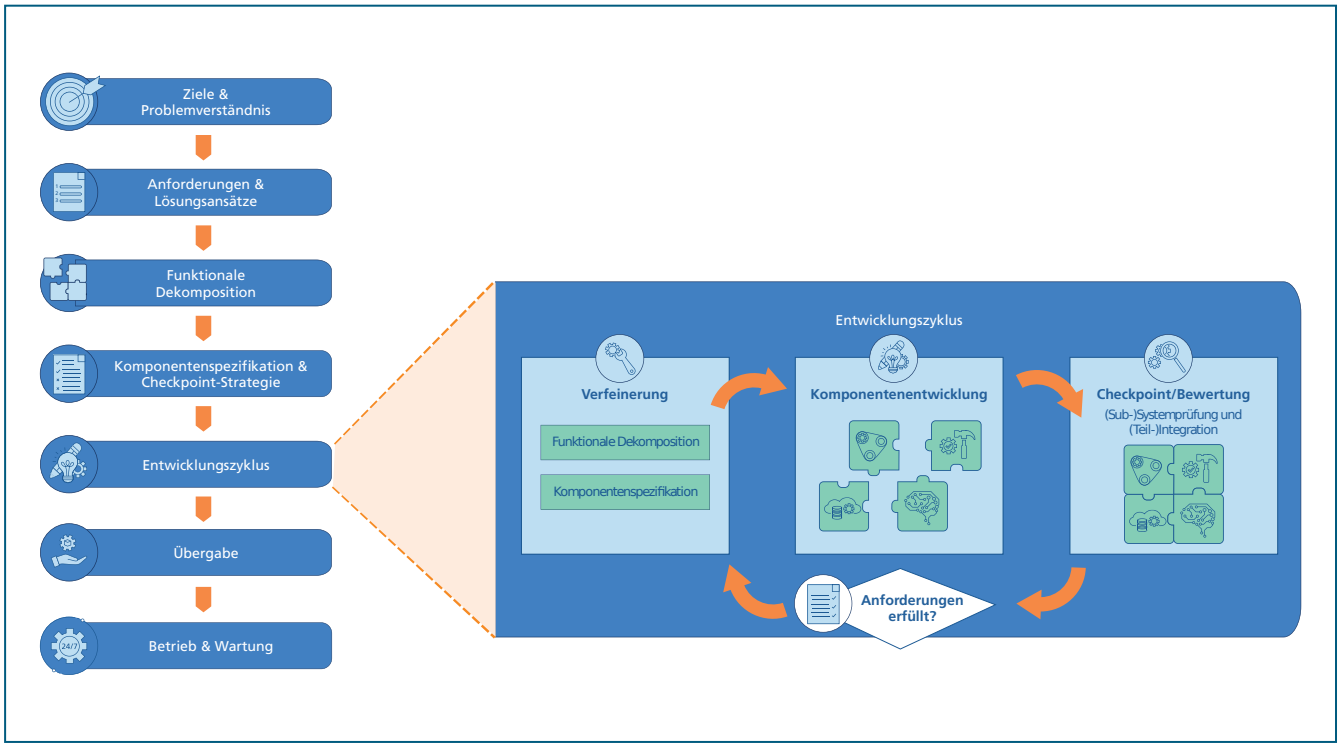


Bild 3: Phasen des Vorgehensmodells PAISE®

7 Ausblick

KI-Engineering ist eine aufkommende Ingenieursdisziplin, die die Methoden, Vorgehensweisen und Technologien der klassischen Ingenieursdisziplinen (u.a. Maschinenbau, Elektrotechnik, Chemieingenieurwesen), mit der Datenwissenschaft (Data Science) sowie der Informatik verbindet. Das vorliegende Whitepaper beschreibt KI-Engineering mit seinen heute sichtbaren wesentlichen Facetten. Die Dynamik im KI-Bereich, in der Standardisierung, in Datenrauminitiativen sowie die europäische Gesetzgebung sowie die neuen Trends in der Informationstechnik werden auf KI-Engineering einen wesentlichen Einfluss haben und die Definition dieser Disziplin prägen. Ein paar davon sollen im vorliegenden Ausblick noch beleuchtet werden. Abschließend führt ein Beispiel einer KI-integrierten Produktionsumgebung verschiedene Aspekte des KI-Engineering zusammen.

7.1 Generative KI

Generative Künstliche Intelligenz hat sich zu einem der faszinierendsten Bereiche der KI entwickelt. Bereits heute sind weite Teile der Content-Erzeugung und Software-Entwicklung durch sie revolutioniert worden und sie hat das Potenzial, die Art und Weise, wie wir mit IT-Systemen interagieren, in vielen weiteren Bereichen entscheidend zu verändern. Der Einsatz von GPT-Modellen, also vortrainierte, generative Transformer-Modelle (generative pretrained transformers) wie ChatGPT oder Stable Diffusion und deren Anwendung in verschiedenen Branchenszenarien, haben eine enorm hohe mediale Aufmerksamkeit erzeugt. Generative Modelle können nicht nur Sprache verstehen und analysieren, sondern sind auch in der Lage, multimediale Inhalte zu generieren, z.B. natürlichsprachliche Texte, Musikstücke oder Programmcode.

Große Sprachmodelle können dazu beitragen, natürlichere und intuitivere Benutzerschnittstellen für KI-Systeme oder allgemein IT-Systeme zu entwickeln. Durch Sprachsteuerung können Benutzer und Benutzerinnen leichter mit KI-Anwendungen interagieren. Sie können, ggf. in der jeweiligen Muttersprache, natürlichsprachliche Anweisungen erteilen, Daten abrufen oder Prozesse steuern, ohne spezielle Kenntnisse in der Bedienung der Systeme bzw. Maschinen zu benötigen. Zudem sind generative Sprachmodell sehr versiert in der Übersetzung natürlicher Sprache in Operationsbefehle technischer Geräte wie z.B. eine Maschine, eine Anlage oder einfach ein Sensor. Liegt beispielsweise eine gut dokumentierte REST-Schnittstelle vor, so sind die Modelle heute schon in der Lage, natürlichsprachliche Anfragen an das Gerät über die REST-Schnittstelle zu erzeugen, über Add-Ons diese auszuführen und auszuwerten.

Über diese neuen Arten der Interaktion hinaus ist der Einsatz generativer KI-Verfahren zukünftig auch für Anwendungen im Kontext des KI-Engineering vorstellbar. Diese haben als Eingangsgrößen jedoch hauptsächlich numerische Daten, wie Sensormesswerte für Maschinen und Anlagen. Vergleichbare Datenmengen wie bei Sprachmodellen liegen (noch) nicht vor. In vielen Einsatzgebieten ist es sogar schlichtweg wirtschaftlich unmöglich, genug Daten durch Sammeln oder Messen zu gewinnen. Für die Bilderkennung, unter echten Einsatzbedingungen, müsste man verschiedenste Umgebungen unter diversen Licht- und Wetterverhältnissen und mit immens vielen Kombinationen filmen. Synthetisch erzeugte Daten durch generative Modelle können hier zukünftig Abhilfe schaffen und der Entwicklung leistungsstarker KI-Anwendungen weiteren Schub geben [28].

Weiteres Potenzial bietet die Kombination von Agentensystemen mit generativer KI. Agentensysteme sind autonome, handlungsfähige Einheiten, die in einer bestimmten Umgebung agieren, Entscheidungen treffen, auf Veränderungen reagieren und so unterschiedliche Konfigurationen testen können. Die generative KI hat die Aufgabe, neue Inhalte oder Informationen in Interaktion mit den Agenten zu erzeugen. Die Integration dieser beiden Technologien kann zukünftig auch im industriellen Umfeld innovative Anwendungen und Einsatzszenarien ermöglichen, sofern es gelingt, die in diesem Bereich vorherrschenden qualitativen Anforderungen zur Verlässlichkeit und

Prognostizierbarkeit der Leistung zu erfüllen. Deshalb ist hierbei KI-Engineering wesentlich.

7.2 Datenräume

Ein Großteil des Aufwands im KI-Engineering fließt in die PAISE® Phase der Datenbereitstellung mit den Teilaspekten der Datenbeschaffung für Experimente und Zielsystem sowie der Datenbewertung gemäß definierten Zielmetriken. Durch die zunehmende Vernetzung der Anlagen und Maschinen in der Industrie 4.0 spielen hierbei die Fragen der Datenherkunft und der Datennutzungsrechte eine wesentliche Rolle. Letztlich ist es die Frage, aus welchem Datenraum (dataspace) die Daten kommen und welche Regeln (zusammengefasst als policies) in dem jeweiligen Datenraum gelten. Als Datenraum wird hier nicht der physische Ort der Datenspeicherung verstanden, also z.B. im Unternehmen (on premise) oder in einer Cloud. Ein Datenraum bezeichnet vielmehr ein Datenintegrationskonzept oder, etwas konkreter, ein datenbasiertes Konzept zur Zusammenarbeit zwischen den Teilnehmern nach einheitlich definierten Regeln unter Wahrung der Datensouveränität. Damit sind im Wesentlichen folgende Zusicherungen verbunden:

- Datenzugriffskontrolle (data access control): Nur autorisierte Teilnehmer dürfen auf Datenbestände nach definierten Zugriffsregeln zugreifen.
- Datennutzungskontrolle (data usage control): Daten dürfen nur für den vom Datenlieferanten intendierten und erlaubten Zweck benutzt werden.
- Datenherkunftsverfolgung (data provenance tracking): Der Datenkonsument muss erkennen können, woher die Daten stammen und ob er sie legal benutzen darf.

Die Einhaltung der Regeln muss durch eine technische Infrastruktur unterstützt und durch juristische Maßnahmen ergänzt werden, so dass das kontrollierte Teilen von Daten auch möglich ist zwischen Unternehmen, die bislang keine bilaterale Vertrauens- oder Geschäftsbeziehung aufgebaut haben. Durch Datenräume können Datenbestände einfacher und effizienter für das Training von KI-Komponenten bereitgestellt werden. Drei Ebenen von Datenräumen sind zu unterscheiden [29]:

1. Zusammenarbeit von Komponenten innerhalb einer Maschine und der IT-Komponenten der Maschine selbst. Ziele u. a. Konfigurierung, Überwachung, Steuerung und Optimierung der Maschine und Komponenten oder Dienste zur vorbeugenden Wartung.
2. Zusammenarbeit von Maschinen in einer Produktionsanlage und der IT-Komponenten der Produktionsanlage (z.B. ein Produktionsleitsystem). Ziele u. a. Überwachung, Steuerung und Optimierung der Produktion, Dienste zur Produktqualitätsprognose oder Energieoptimierung.
3. Zusammenarbeit von Produktionsanlagen über Fabrik- und Unternehmensgrenzen hinaus. Ziele u.a. Lieferkettenmanagement, Nachhaltigkeitsberechnungen (z. B. CO_2-Fußabdruck), Dienste zur Resilienzsteigerung und Flexibilisierung

Datenräume können zudem durch Dienste zur Datenaufbereitung angereichert werden. Hilfreich für das KI-Engineering ist z. B. die Aufbereitung von Features für Modelle, etwa durch Feature Stores. Datenräume können dabei helfen, indem sie Funktionen wie Dateikonvertierung, -komprimierung und -bereinigung, die für die Vorbereitung der Daten erforderlich sind, unterstützen. Schließlich können KI-Systeme aufgrund ihrer Fähigkeit, Erkenntnisse aus Daten zu gewinnen, sehr wertvoll sein. Datenräume bieten hier eine sichere Möglichkeit, in Lieferbeziehungen zwischen Unternehmen die Vertraulichkeit und Integrität der Daten sicherzustellen und unbefugten Zugriff zu verhindern. Die vom BMWK geförderte Datenrauminitiative Manufacturing-X entwickelt hierfür die notwendige Infrastruktur und validiert sie an ausgewählten Anwendungsfällen [30].

7.3 Föderiertes Lernen

Föderiertes Lernen (Federated Learning) ist eine Methode des maschinellen Lernens, bei der Modelle auf dezentralen Geräten trainiert werden, ohne dass die Daten von diesen Geräten ausgetauscht werden müssen.

Für die KI-basierte Optimierung unserer Produktionsanlage in Schwabmünchen reicht es nicht, auf unsere eigenen Sensordaten zuzugreifen. Wir müssen auch Qualitätsdaten der Vorprodukte (u.a. Wolframpulver) mit einbeziehen und geben auch Daten weiter an unsere Kunden im Kontext der gesamten Wertschöpfungskette.«

Dipl.-Ing. Ingo Hild, Plant Manager, ams OSRAM Group

Stattdessen werden nur die Modellparameter ausgetauscht und in einem globalen Modell aggregiert. Insbesondere in Anwendungsfällen, in denen Privatsphäre, Datensicherheit und -hoheit eine Rolle spielen, ist diese Art der Modellerstellung vorteilhaft. Besondere Vorteile ergeben sich, wenn auf Datenräume, wie oben beschrieben, zurückgegriffen werden kann [31]. Dadurch können Daten von verschiedenen Quellen und Standorten nach einheitlich definierten Regeln verfügbar gemacht werden können. Dies erhöht zum einen die Generalisierbarkeit der einzelnen Modelle, was zu einer verbesserten Modellqualität führt. Zum anderen erhöht sich die Skalierbarkeit der Modelle, da mehr Datenquellen zur Verfügung stehen.

Insbesondere im Maschinenbau besteht großes Potenzial zur Nutzung von Methoden des föderierten Lernens, da Hersteller intelligente Services für ausgelieferte Maschinen anbieten können, ohne dafür die Rohdaten ihrer Kunden teilen zu müssen. Beispielsweise kann man sich einen Hersteller vorstellen, der seine Maschine oder Anlage mit einem initialen ML-Modell ausliefert, welches mit Hilfe der Daten aus dem Betrieb anschließend optimal an den Anwendungszweck angepasst wird. Dieses Modell wiederum kann dem Hersteller zur Verfügung gestellt werden, der damit die Gesamtqualität seiner ML-Modelle erhöhen kann.

7.4 Erklärbare und vertrauenswürdige KI

Erklärbare und vertrauenswürdige KI kann beim KI-Engineering auf verschiedene Weise unterstützen und wird an einigen Stellen auch vorausgesetzt. Die verfügbaren Ansätze zur Erklärbarkeit reichen aktuell noch nicht aus, um den Einsatz von KI-Verfahren für alle Anwendungsfälle zu ermöglichen. Wünschenswert ist beispielsweise eine Erklärbarkeit auf semantischer bzw. symbolischer Ebene, welche für Maschinenbediener ein intuitives Verständnis von Fehlermodi oder Ergebnissen liefert.

Ferner müssen bestehende Safety- und Security-Prüfverfahren und entsprechende Standards dahingehend erweitert werden, dass die Zulassung von KI-Komponenten und der Gesamtsysteme nach klaren Kriterien durchgeführt werden kann. Erklärbarkeit und Transparenz sind auch hier integraler Bestandteil für entsprechende Sicherheitsargumentationen.

Zudem werden neue, auf KI-Verfahren und KI-Engineering zugeschnittene Standards benötigt, um generalisierte erweiterte Methoden zur Erklärbarkeit und Vertrauenswürdigkeit nach einheitlichen Grundkonzepten entwickeln zu können. Die zweite Ausgabe der Deutschen Normungsroadmap KI [6] enthält KI-Engineering als eigenständiges Kapitel und schlägt u.a. folgenden Normungsbedarf vor:

- Modellierung von KI-Systemen mit systemtechnischen und anwendungsbezogenen Aspekten.
- Metabeschreibung von KI-Verfahren.
- Metadatenbeschreibungen von Eingangs- und Ausgangsdatensätzen von ML-Verfahren.
- Taxonomie, textuelle und ggf. formale Beschreibung von Qualitätskriterien KI-basierter Systeme u. a. Verlässlichkeit, Zuverlässigkeit, Planbarkeit, Kontrollierbarkeit, …
- Metriken zur Erklärbarkeit mit dem Ziel, einen Kompromiss zwischen Erklärbarkeit der angewandten maschinellen Lernverfahren (Gedankenmodell der Anwender) und der Genauigkeit bzw. Güte der Erklärungen herstellen zu können.
- Nutzung semantischer Modelle in der Erklärbarkeit.
- Eine wichtige Rolle wird in diesem Zusammenhang den aktuell entstehenden harmonisierten Normen zur Operationalisierung der europäischen KI-Verordnung [22] zukommen. Das bestehende Normenwerk beinhaltet aktuell noch stellenweise Vorgaben, welche den Einsatz von KI-Verfahren behindern. Deshalb laufen aktuell Projekte des BMWK und DIN, die existierende Normen auf deren KI-Tauglichkeit überprüfen [32].

7.5 KI-integrierte Produktionsumgebung

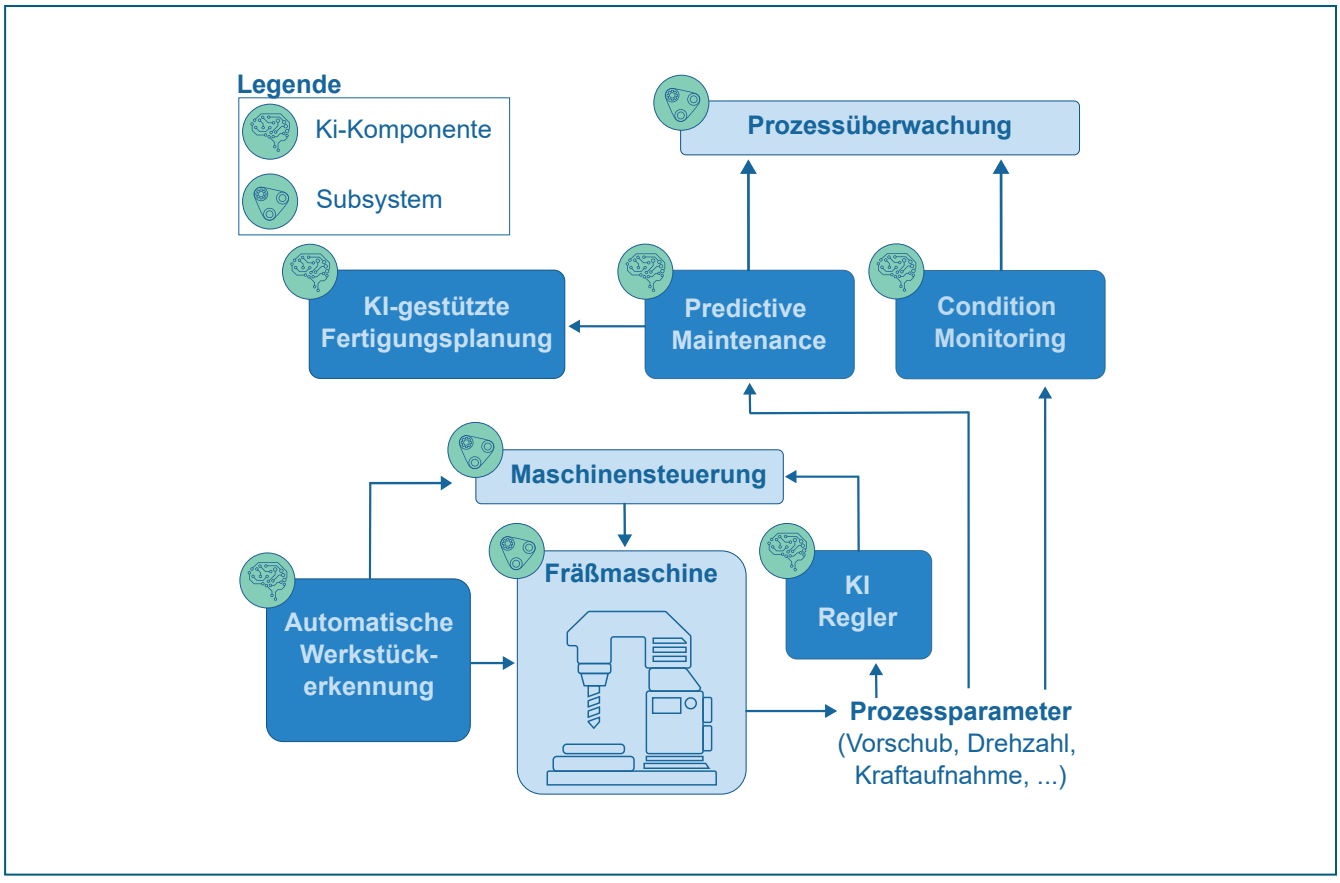
Den großen Vorteil von KI-Engineering erkennt man, wenn man nicht nur die

Integration von KI-Komponenten in einer einzelnen Anwendung betrachtet, sondern ein Gesamtsystem, bei dem viele einzelne KI-Subsysteme und »klassische« Subsysteme zum Einsatz kommen, die sich untereinander beeinflussen, voneinander abhängen und integriert werden müssen. Hier ist es umso wichtiger, einen ganzheitlichen Ansatz bei der Entwicklung zu wählen. Ein solcher Anwendungsfall mit mehreren Subsystemen und KI-Komponenten ist im Bild 4 unten am Beispiel einer Fräsmaschine dargestellt.

Die Fräsmaschine erkennt zu Beginn des Bearbeitungsprozesses in der KI-gestützten Werkzeugerkennung das ihr übergebene Werkstück automatisch und übergibt die Werkstückinformation an die Maschinensteuerung, welche die Fräsmaschine entsprechend parametrisiert. Während das Werkstück bearbeitet wird, übernimmt ein KI-Regler die Feinjustierung der Prozessparameter, um ein optimales Bearbeitungsergebnis zu erzielen.

Parallel dazu überwacht eine KI-basierte Condition Monitoring-Einheit den Prozess und gibt eine Warnung aus, falls kritische Abweichungen erkannt werden. Außerdem prognostiziert ein KI-basierte Predictive Maintenance-Algorithmus anhand der Werkzeugabnutzung den nächsten Fräsertausch und übergibt diese Information dem KI-gestützten Produktionsplanungssystem, welches dies in der Anlagenverfügbarkeit berücksichtigt.

Bild 4: Anwendungsbeispiel einer KI-integrierten Produktionsumgebung



8 Relevante Projekte und Initiativen

8.1 ML4P – Machine Learning for Production

Im Rahmen des Fraunhofer-Leitprojekts »ML4P – Machine Learning 4 Production« [33] (2018 –2022) wurde das Ziel verfolgt, ein werkzeuggestütztes Vorgehensmodell für die Entwicklung von ML-Verfahren in der Produktion zu entwerfen. ML4P sollte weitreichende Optimierungsmöglichkeiten in industriellen Produktionsprozessen durch den Einsatz von maßgeschneiderten ML-Methoden und geeigneten Softwaretools systematisch identifizieren und gezielt nutzen.

Mit dem zunehmenden Einsatz von ML-Methoden in der industriellen Praxis zeigte sich, dass bei der Umsetzung ein besonderes Augenmerk auf das Wissensmanagement, die Organisation der Projektteams und die Standardisierung der eingesetzten Werkzeuge gelegt werden muss. In kleineren Umsetzungsprojekten konnten die Beteiligten bisher sowohl das Wissen über die Anwendungsdomäne als auch die theoretischen Grundlagen der ML-Methoden vollständig verstehen. Bei umfangreichen Projekten ist dies jedoch aufgrund ihrer Komplexität nicht mehr möglich. Daher erfordert die effiziente Durchführung von ML-Projekten eine klare Strukturierung der Aufgaben und Verantwortlichkeiten.
Ein wesentliches Ergebnis der Forschungsarbeiten in ML4P ist die Ausarbeitung des ML4P-Vorgehensmodells. Es strukturiert die notwendigen Arbeitsschritte – Informationsgewinnung, Analyse, Lernen und Betrieb – für den effektiven Einsatz von ML. Die Festlegung spezifischer Rollen und die Einführung der »Prozessakte« als zentrales Dokument tragen wesentlich zur erfolgreichen Integration von ML-Methoden in die industrielle Praxis bei. Für die software-seitige Umsetzung von industrietauglichen ML-Lösungen wurden in ML4P zahlreiche Softwarewerkzeuge entwickelt. Diese decken die typischen Arbeitsaufgaben – von der Datenerfassung über die interaktive Analyse und das Modelllernen bis hin zum dauerhaften Betrieb – ab.

Ein Beispiel für die Nutzung der ML4P-Technologien in der Praxis sind robuste Lösungen für die Warmumformung von Blech, das Biegen von Glas und die Produktion von Membranfiltern. Die Erfahrungen, die in Zusammenarbeit mit Industriepartnern gesammelt wurden, haben gezeigt, dass mit ML-Methoden Optimierungspotenziale in industriellen Produktionsprozessen systematisch erschlossen werden können. Das ML4P Vorgehensmodell war eine wesentliche Grundlage für die Entwicklung des umfassenderen KI-Engineering Vorgehensmodell PAISE® (vgl. Kapitel 6.5).

8.2 CC-KING – Kompetenzzentrum für KI-Engineering

Um KI-Engineering als Methodik genauer zu fassen und als eigenständige Ingenieursdisziplin voranzutreiben, hat sich in der TechnologieRegion Karlsruhe seit August 2020 das Kompetenzzentrum für KI-Engineering CC-KING gebildet, gefördert durch das Ministerium für Wirtschaft, Arbeit und Tourismus Baden-Württemberg [34].

CC-KING ist ein Konsortialvorhaben des Fraunhofer IOSB, Karlsruhe, mit dem FZI Forschungszentrum Informatik und dem Karlsruher Institut für Technologie (KIT). Es wird geleitet vom Fraunhofer IOSB. CC-KING erarbeitet die wissenschaftlichen Grundlagen und Methoden für KI-Engineering, entwickelt Software-Werkzeuge zu deren Unterstützung und Anwendung und demonstriert die Ergebnisse anhand von praxisnahen Anwendungsfällen vor allem aus den Domänen industrielle Produktion und Mobilität.

Die wissenschaftliche Befassung mit KI-Engineering und vor allem die Maßnahmen zur Umsetzung in der Industrie und im Mittelstand haben gezeigt, dass die Entwicklung der Branche hier am Anfang einer längeren Wegstrecke steht. Auf der wissenschaftlichen Seite ist festzustellen, dass es zahlreiche punktuelle Ansätze zur Lösung dieser Herausforderungen gibt, hingegen noch keine in sich schlüssige methodische Gesamtdarstellung existiert. Hier

konnte CC-KING mit PAISE® einen Mehrwert erzielen und ein weltweit erstes Vorgehensmodell für systemorientiertes KI-Engineering vorlegen [35].

Um den Technologietransfer für KI-Engineering auch über den Förderzeitraum von CC-KING weiterzuführen, wurde ein Forum KI-Engineering eingerichtet, das verschiedene Maßnahmen und assoziierte Projekte bündelt. CC-KING bleibt als Kompetenzzentrum der Kern des Forums und ist zuständig für die inhaltliche und fachliche Weiterentwicklung der Methodik KI-Engineering hin zu einer Disziplin KI-Engineering.

8.3 KI-Lernlabor – Künstliche Intelligenz für den Mittelstand

Viele Unternehmen haben bereits das Potenzial von KI-Anwendungen für sich erkannt. Konkrete Umsetzungen, d.h. die Ausleitung durchgeführter Use Case Studien bis in die Produktphase, haben jedoch in der großen Breite der Wirtschaft bisher nur wenig stattgefunden. Dies ist vor allem in der Unkenntnis über die Möglichkeiten von KI, einer allgemeinen Verunsicherung im Umgang mit KI-Anwendungen, und einem Mangel an KI-Experten aufgrund fehlender Aus- und Weiterbildung für KI-Experten begründet. Hinzu kommen außerdem technische Herausforderungen durch neuartige Infrastrukturen (u.a. Internet of Things, Edge / Cloud-Plattformen) und Systemarchitekturen (u.a. verteilte Big Data Systeme) sowie ein hochdynamisches Ökosystem von KI-Software-Frameworks und KI-Verfahren.

Ziel des vom BMBF geförderten Projektes »KI-Lernlabor« [36] war es, Formate zu entwickeln, um insbesondere kleine und mittelständische Unternehmen (KMU) aller Branchen unmittelbar an aktuelle Forschungsthemen der KI heranzuführen und diese daran praktisch teilhaben zu lassen. Forschung auf dem Bereich KI soll für die beteiligten Unternehmen und deren Mitarbeitenden fühl- und erlebbar gemacht werden. Dies wurde durch die Etablierung eines KI-Lernlabors ermöglicht. In diesem werden KMU konkret bei der Ideenfindung über die Einführung und Umsetzung bis hin zum Betrieb von KI-Anwendungen mit diversen Maßnahmen begleitet. Wir möchten somit insbesondere KMU für das Thema KI begeistern, qualifizieren und in der Praxis bei der Umsetzung unterstützen. Das KI-Lernlabor wurde dafür im Rahmen des Projektes auf dem Fraunhofer-Campus Schloss Birlinghoven als auch in Form von virtuellen Veranstaltungen unter Beteiligung verschiedener Fraunhofer-Einrichtungen etabliert.

Neben verschiedenen Schulungsformaten und Co-Working-Formaten wurde hierbei u.a. auch der Prozessansatz MLOps (vgl. Kapitel 6.3) evaluiert und mittels eines Whitepapers für die Weiterbildung und den Know-how-Transfer in die Unternehmen konzipiert und angeboten.

8.4 KI-NRW Studie zum Einsatz von MLOps

Die Kompetenzplattform Künstliche Intelligenz Nordrhein-Westfalen (KI.NRW) und das Fraunhofer IAIS führten bis Sommer 2023 eine Studie zum Thema MLOps (vgl. Kapitel 6.3) durch mit dem Ziel, Informationen über den Stand der Entwicklung und den Bedarf an Unterstützung zu ermitteln, den Unternehmen im Bereich MLOps haben.

Im Rahmen der Durchführung der Studie tauschten sich MLOps Experten des Fraunhofer IAIS direkt mit den relevanten Ansprechpartnern aus den Unternehmen aus. Zur Aufnahme der Informationen wurden mehr als 25 qualifizierte Interviews mit Firmen durchgeführt. Die Auswahl der Firmen erfolgte über verschiedene Branchen und reicht von kleineren Firmen inkl. Start-Ups bis hin zu Konzernen.

Die Fragen in den Interviews richteten sich nach den Phasen des MLOps Zyklus. Zu jeder Phase des Zyklus werden mehrere Detailfragen zum Umfang der Aktivitäten in den jeweiligen Firmen gestellt. Die Ergebnisse der Interviews werden im Sommer 2023 ausgewertet und sollen noch im Jahr 2023 in anonymisierter Form als Studie veröffentlicht werden.

Bei Materna verstehen wir die immense transformative Kraft der Data Economy mit Künstlicher Intelligenz. Durch unsere Expertise in der Entwicklung und Implementierung von Datenräumen mit fortschrittlichen Machine Learning- und Analyse-Werkzeugen und der Anwendung von MLOps-Praktiken ermöglichen wir dem öffentlichen Sektor und Unternehmen effizient vom Prototypen in die Anwendung unter Echtzeitbedingungen zu kommen.«

Thomas Feld, VP Data Economics, Materna SE

8.5 KI-Allianz Baden-Württemberg eG – Teilvorhaben Datenplattform

Die KI-Allianz Baden-Württemberg eG verfolgt das Ziel, einen Innovationsraum für KI-Anwendungen zu schaffen mit integrierten, branchenübergreifenden und anwendungsorientierten Datenräumen, KI-Reallaboren und Testfeldern für Unternehmen, Start-ups und Wissenschaft. Dafür haben sich zahlreiche Gebietskörperschaften, Regionen und Städte in Baden-Württemberg zu einer eingetragenen Genossenschaft zusammengeschlossen [37]. Unternehmen, Verbände und Forschungseinrichtungen können sich dieser Genossenschaft anschließen.
Zudem laufen im Kontext der KI-Allianz BW verschiedene Entwicklungs- und Anwendungsvorhaben zur Förderung und Unterstützung von KI-Anwendungen. Ein erstes Teilvorhaben ist die Konzeption und der prototypische Aufbau einer branchenübergreifenden KI-bezogenen Datenplattform, um KI-Anwendungen mittels eines offenen Datenraums (s.o.) die notwendige Infrastruktur bereitzustellen. Das Teilvorhaben Datenplattform wird gefördert vom Ministerium für Wirtschaft, Arbeit und Tourismus Baden-Württemberg und läuft von Mitte 2023 bis Ende 2025 unter Leitung des Fraunhofer IOSB. Es adressiert insbesondere die Anwendungsdomänen Produktion, Mobilität, Smart City und Gesundheit [38].

8.6 Fraunhofer Angebote

- Schulungs- und Weiterbildungsangebote der Fraunhofer-Gesellschaft zu Data Science und KI-Verfahren finden Sie auf der Web-Seite der Fraunhofer-Allianz Big Data und Künstliche Intelligenz unter https://www.bigdata-ai.fraunhofer.de/.
- Beratungsangebote und Praxisbeispiele der Fraunhofer-Gesellschaft zu Informed Machine Learning und zum Einsatz von KI von der Edge bis in die Cloud finden Sie unter https://www.cit.fraunhofer.de
- Beratungs- und Schulungsangebote der Fraunhofer-Gesellschaft und Praxisbeispiele zu KI-Engineering finden Sie auf der Web-Seite des Kompetenzzentrums für KI-Engineering Karlsruhe (CC-KING) unter https://www.ki-engineering.eu/.

9 Referenzen

[1] Deloitte, »Fueling the AI transformation, https://www2.deloitte.com/de/de/pages/trends/ki-studie-2022.html,« 2022.

[2] Bitkom Research, »Künstliche Intelligenz – Wo steht die deutsche WIrtschaft?,« 2022.

[3] VDI, 2022. [Online]. Available: https://www.vdi.de/news/detail/einsatz-von-ki-die-erwartungen-erfuellen-sich-noch-nicht.

[4] J. Pfrommer, T. Usländer und J. Beyerer, »KI-Engineering – AI Systems Engineering: Systematic development of AI as part of systems that master complex tasks,« Automatisierungstechnik, vol. 70, no. 9, https://doi.org/10.1515/auto-2022-0076, pp. 756-766, 2022.

[5] ISO/IEC 22989:2022, »Information technology – Artificial intelligence – Artificial intelligence concepts and terminology,« [Online]. Available: https://www.iso.org/standard/74296.html.

[6] DIN, »Deutsche Normungsroadmap Künstliche Intelligenz Version 2,« 2022. [Online]. Available: https://www.din.de/de/forschung-und-innovation/themen/kuenstliche-intelligenz/fahrplan-festlegen.

[7] IDTA (Industrial Digital Twin Association), »Specification of the Asset Administration Shell – Part 1: Metamodel,« 2023. [Online]. Available: https://industrialdigitaltwin.org/wp-content/uploads/2023/06/IDTA-01001-3-0_SpecificationAssetAdministrationShell_Part1_Metamodel.pdf.

[8] »Informed Machine Learning – Technologienkern Machinelles Lernen des CCIT,« [Online]. Available: https://www.cit.fraunhofer.de/de/Technologiekerne/maschinelles-lernen.html.

[9] Plattform Industrie 4.0, »Technologieszenario »Künstliche Intelligenz in der Industrie 4.0«,« Bundesministerium für Wirtschaft und Energie (BMWi), Berlin, 2019.

[10] A. Mayr, B. Lutz, M. Weigelt, T. Gläßel, J. Seefried, D. Kißkalt und J. Franke, »Elektromotorenproduktion 4.0: Potenziale des maschinellen Lernens in der Elektromotorenproduktion am Beispiel des Laserschweißens von Hairpins,« Zeitschrift für wirtschaftlichen Fabrikbetrieb, vol. 114, no. 3, pp. 145-149, 2019.

[11] »Mantelflächenprüfung mit NELA VisionCheck,« [Online]. Available: https://www.nela.de/de/produkt/anwendungsbeispiel-oberflaechenpruefung-von-muttern-und-scheiben. [Zugriff am 30 06 2023].

[12] »TableSort® – Schüttgutsortierer »to go«,« [Online]. Available: https://www.iosb.fraunhofer.de/de/projekte-produkte/tablesort.html.

[13] Fraunhofer IOSB, »visIT Themenheft Recycling,« 04/2017. [Online]. Available: https://www.iosb.fraunhofer.de/content/dam/iosb/iosbtest/documents/publikationen/visit/visIT%20Recycling.pdf.

[14] Fraunhofer IOSB, »visIT, Erklärbare KI im Einsatz,« 2023. [Online]. Available: https://services.iosb.fraunhofer.de/visIT/xai/#0.

[15] »ACME 4.0 – Akustische Zustandsüberwachung für Industrie 4.0,« [Online]. Available: https://www.idmt.fraunhofer.de/de/institute/projects-products/projects/acme.html. [Zugriff am 30 06 2023].

[16] »Condition Monitoring für Laser,« [Online]. Available: https://www.trumpf.com/de_DE/produkte/services/services-fuer-laser/monitoring-analyse/condition-monitoring-fuer-laser/. [Zugriff am 30 06 2023].

[17] »SCHUNK startet in eine neue Ära der Werkzeugspannung,« [Online]. Available: https://www.hannovermesse.de/de/news/news-fachartikel/schunk-startet-in-eine-neue-aera-der-werkzeugspannung.

[18] »GT Auto Tuner,« [Online]. Available: https://www.siemens-energy.com/global/en/offerings/services/digital-services/gt-autotuner.html.

[19] »KI-basierte Fertigungszeitenprognose bei Sykatec,« [Online]. Available: https://www.ipa.fraunhofer.de/de/referenzprojekte/ki_basierte_fertigungszeitenprognose_bei_sykatec.html.

[20] A. Schmitz, M. Akila, D. Hecker, M. Poretschkin und S. Wrobel, »Schmitz, Anna, et al. »The why and how of trustworthy AI: An approach for systematic quality assurance when working with ML components,« at-Automatisierungstechnik 70.9, pp. 793-804, 2022.

[21] HEG-KI (Hochrangige Expertengruppe für künstliche Intelligenz), »Ethik-Leitlinien für eine vertrauenswürdige KI,« Europäische Kommission, 2019.

[22] European Commission, »Proposal for a Regulation laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts,« Brüssel, 2021.

[23] M. Poretschkin (Ed.), »Leitfaden zur Gestaltung vertrauenswürdiger Künstlicher Intelligenz – KI-Prüfkatalog,« 2021. [Online]. Available: https://www.iais.fraunhofer.de/de/forschung/kuenstliche-intelligenz/ki-pruefkatalog.html.

[24] P. Chapman, J. Clinton, R. Kerber, T. Khabaza, T. Reinartz, C. Shearer und R. Wirth, »CRISP-DM 1.0: Step-by-step data mining guide,« SPSS inc, 9(13), 2000.

[25] IBM, »ASUM-DM: Analytics Solutions Unified Method,« 2016. [Online]. Available: https://public.dhe.ibm.com/software/data/sw-library/services/ASUM.pdf.

[26] VDI/VDE 2206, »Entwicklung mechatronischer und cyber-physischer Systeme,« 2021.

[27] C. Hasterok und J. Stompe, »PAISE® – Process model for AI systems engineering,« Automatisierungstechnik, vol. 70, no. 9, https://doi.org/10.1515/auto-2022-0020, pp. 777-786, 2022.

[28] D. Hecker, A. Voss, G. Paaß und T. Wirtz, »Big Data 2.0–mit synthetischen Daten KI-Systeme stärken,« Wirtschaftsinformatik & Management, 15(2), pp. 161-167, 2023.

[29] T. Usländer, »KI-Engineering für die Produktion im Kontext von Datenräumen,« de Gruyter, Zeitschrift für wirtschaftlichen Fabrikbetrieb (zwf) vol. 118, no. 5, 2023.
[30] B. Otto, J. Seidelmann, J. Schmeltung und O. Sauer, »Bauplanstudie: Datenraum Manufacturing-X,« 2023. [Online]. Available: https://www.zvei.org/fileadmin/user_upload/Presse_und_Medien/Pressebereich/2023_059_Manufacturing_X/Vorstudie_Datenraum_Manufacturing-X_ZVEI-VDMA-Fraunhofer.pdf.
[31] D. Hecker, A. Voss und S. Wrobel, »Data Ecosystems: A New Dimension of Value Creation Using AI and Machine Learning,« in Designing Data Spaces, Springer International Publishing, 2022, pp. 211-224.
[32] DIN, »KI Tauglichkeit von Normen,« [Online]. Available: https://www.din.de/de/forschung-und-innovation/themen/kuenstliche-intelligenz/projekte-zu-ki-und-normung/ki-tauglichkeit-von-normen/ki-tauglichkeit-von-normen-872324.
[33] Fraunhofer IOSB, »ML4P Vorgehensmodell - White Paper,« [Online]. Available: https://www.iosb.fraunhofer.de/content/dam/iosb/iosbtest/documents/projekte/ml4p/ML4P_whitepaper.pdf.
[34] CC-KING, »Kompetenzzentrum für KI-Engineering Karlsruhe,« Fraunhofer IOSB, [Online]. Available: https://www.ki-engineering.eu/.
[35] CC-KING, »PAISE: Das KI-Engineering Vorgehensmodell,« [Online]. Available: https://www.ki-engineering.eu/de/wissen-tools/paise.html.
[36] A. Zimmermann, D. Wegener, K.-H. Sylla, C. Martens und N. Beck, »Zukunftssichere Lösungen für maschinelles Lernen - Machine Learning operations (MLOps) – Prozesse für Entwicklung, Integration und Betrieb,« Fraunhofer-Institut für Intelligente Analyse- und Informationssysteme IAIS, Schloss Birlinghoven, Sankt Augustin, 2020.
[37] KI-Allianz Baden-Württemberg eG, »KI-Allianz BW,« [Online]. Available: https://www.ki-allianzbw.de/.
[38] Fraunhofer IOSB, »KI-Allianz BW: Startschuss für Aufbau einer KI-Datenplattform für die Wirtschaft,« [Online]. Available: https://www.iosb.fraunhofer.de/de/presse/presseinformationen/2023/ki-allianz-bw-datenplattform-start.html.
[39] S. Huber, H. Wiemer, D. Schneider und S. Ihlenfeldt, »DMME: Data mining methodology for engineering applications – a holistic extension to the CRISP-DM model,« Procedia CIRP, Vol. 79, pp. 403-408, 2019.

Um Methoden der Künstlichen Intelligenz (KI) in IT-Systemen der industriellen Produktion nachhaltig und operativ einzusetzen, bedarf es der Methodik des KI-Engineering. KI-Engineering adressiert die systematische Entwicklung und den Betrieb von KI-basierten Lösungen als Teil von Systemen, die komplexe Aufgaben erfüllen. Ziel ist es, das Innovations- und Optimierungspotenzial von KI-Verfahren in der industriellen Produktion nutzen zu können.

Das Whitepaper spannt die Dimensionen für KI-Engineering-Anwendungen auf, umreißt die qualitativen Anforderungen in der Entwicklung und im Betrieb unter dem Blickwinkel des Anwenders und Entscheiders. Verschiedene Anwendungsfälle werden in vier Autonomiestufen eingeordnet: von KI-basierten Assistenzfunktionen bis hin zu autonomen und adaptiven Systemen. Zudem werden passende Lösungsmethoden aufgezeigt.

Ein Kapitel widmet sich den technischen und organisatorischen Schulden beim Einsatz von KI-Methoden. Hierin wird als Antwort das KI-Engineering-Vorgehensmodell PAISE® im Kontext bestehender Modelle aus dem Data Mining und dem Software-Engineering erläutert. Im Anschluss werden relevante Initiativen und Projekte beschrieben und anstehende Entwicklungen umrissen.

Fraunhofer IOSB

Fraunhofer IAIS

Fraunhofer VERLAG

Fraunhofer CCIT

ISBN 978-3-8396-1943-8

9 783839 619438