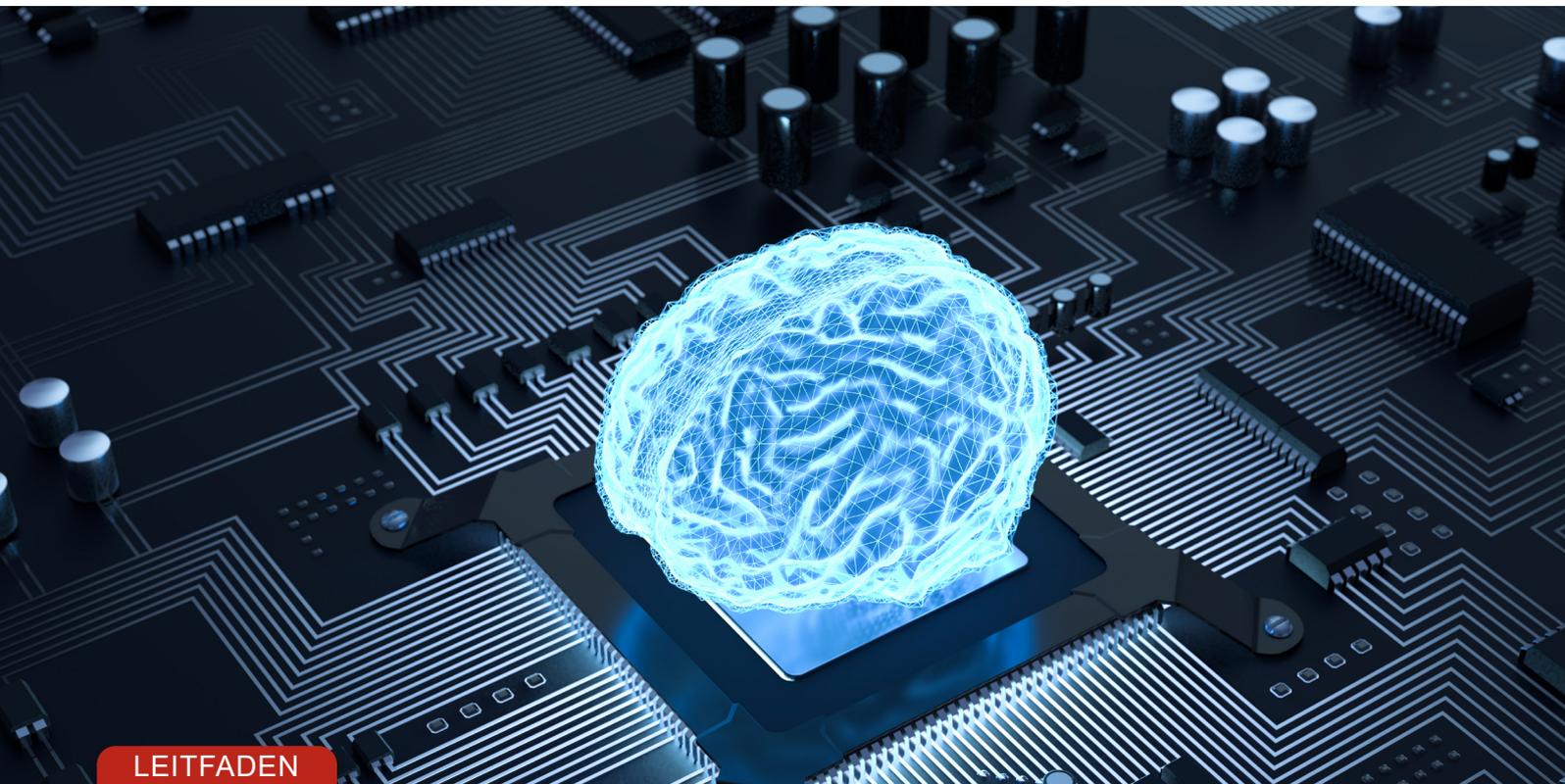




**Mittelstand 4.0**  
Kompetenzentrum  
Stuttgart

#digitalinBW



LEITFADEN

# Machine Learning Architekturen für Entscheidende

Eine Orientierungshilfe zum Einstieg in intelligente Produkte und Dienstleistungen

Mittelstand-  
Digital 

Gefördert durch:



aufgrund eines Beschlusses  
des Deutschen Bundestages

# Inhalt

1. An wen richtet sich der Leitfaden? . . . . .	4
2. Inhalte des Leitfadens und Einführung . . . . .	5
3. Phasen der digitalen Produktentwicklung . . . . .	6
4. Wie funktioniert künstliche Intelligenz? . . . . .	10
5. Maschinelles Lernen - wie geht das? . . . . .	14
6. IT-Architekturen für Maschinelles Lernen. . . . .	20

Ihr Kontakt zu uns . . . . .	.26
Quellenverzeichnis . . . . .	.28
Bildnachweis . . . . .	.29
Impressum. . . . .	.29

# 1. An wen richtet sich der Leitfaden?

Die Integration von Sensorik in Produkte und die Vernetzung dieser Produkte ermöglichen es auch kleinen und mittleren Unternehmen, neue Daten und Informationen zu gewinnen, mit denen sie durch Datenauswertung und Maschinelles Lernen einerseits den Funktionsumfang ihrer Produkte verbessern und andererseits begleitende und intelligente Dienste rund um ihre Produkte anbieten können.

Um solche Lösungen umzusetzen, werden entsprechende IT-Architekturen benötigt, die für die Integration von Geräten, Produkten, Maschinen und anderen Dingen in das Internet der Dinge geeignet sind. Der englische Begriff für das Internet der Dinge ist Internet of Things, oder kurz: IoT.

Der vorliegende Kurzleitfaden wendet sich an Führungskräfte und Entscheidungsträger in Unternehmen, die ihre Produkte und Dienstleistungen weiterentwickeln wollen und dabei die Vorteile und Möglichkeiten der Digitalisierung speziell auch unter Einbindung des Maschinellen Lernens (ML) in Zusammenhang mit vernetzten sensorbasierten Systemen nutzen möchten.

Er bietet eine Übersicht für Entscheider, die sich vor der Umsetzung intelligenter Produkte und Dienstleistungen einen Überblick über die technischen Zusammenhänge und Erfordernisse einer ML-Architektur verschaffen möchten.

## 2. Inhalte des Leitfadens und Einführung

Der Leitfaden setzt ein grundlegendes Verständnis darüber voraus, was unter Smarten Produkten und Dienstleistungen zu verstehen ist. Diese Grundlagen sind im [Leitfaden »Smarte Produkte und Dienstleistungen«](#) des Mittelstand 4.0-Kompetenzzentrums Stuttgart dargestellt. Darüber hinaus baut der hier vorliegende Leitfaden auf dem Leitfaden »IoT Architekturen für Entscheider« auf, in dem die wesentlichen Bestandteile von IoT-Architekturen und die Möglichkeiten und Konsequenzen des In- und Outsourcings von damit verbundenen IT-Aufgaben beschrieben sind.

Der hier vorliegende Leitfaden beschreibt in Ergänzung dazu, wie die Umsetzung von Maschinellern Lernen als Voraussetzung für Anwendungen Künstlicher Intelligenz aussehen kann und welche Erweiterungen man speziell für diesen Zweck an einer IoT-Architektur vornehmen muss. Zu diesem Zweck werden zunächst die verschiedenen »Phasen der digitalen Produktentwicklung« erläutert und welche Mehrwerte maschinelles Lernen und künstliche Intelligenz in die digitale Produkt- und Dienstleistungsentwicklung einbringen können. Hieraus lässt sich ein Vorgehen zur Weiterentwicklung von konventionellen bzw. digitalen Produkten und Services zu intelligenten Produkten und Services ableiten.

Im Anschluss wird die prinzipielle Funktionsweise von künstlicher Intelligenz und maschinellem Lernen erläutert, um im nächsten Schritt dann aufzeigen zu können, welche Architekturbestandteile für die Umsetzung von maschinellem Lernen zu einer IoT-Architektur hinzugefügt werden müssen.

## 3. Phasen der digitalen Produktentwicklung

Die klassische Weiterentwicklung von Produkten dreht sich vor allem um dingbasierte Funktionen, also am Beispiel einer Armbanduhr gesprochen, beispielsweise an der Weiterentwicklung der Robustheit gegen Wasser und Stürze oder der Verbesserung der Genauigkeit, mit der die Zeit angezeigt wird. Demgegenüber steht die IT-basierte Weiterentwicklung, bei der es vor allem darum geht, Daten durch die Vernetzung des Produkts in einem größeren Kontext zur Verfügung zu stellen. Dies geht meist einher mit der Verwendung von Sensorik in den Produkten, über die die Erhebung der Daten ermöglicht wird. Es schließt die Vernetzung zur Weitergabe der Daten ein und schließlich auch die Weiterverarbeitung der Daten außerhalb des eigentlichen Produkts. Arbeiten der Universität St. Gallen haben in diesem Zusammenhang eine Produkt-Service-Logik im Internet der Dinge (engl.: Internet of Things, kurz: IoT) beschrieben, die genau auf das oben beschriebene Prinzip eingeht. Basis für erfolgreiche oder gar disruptive Geschäftsmodelle scheint also die Erhebung von Daten in Produkten und deren Weiterverwendung in global nutzbaren digitalen Services zu sein – weit über die Produktgrenzen oder die Produktumgebung hinaus.

Eine zusätzliche Möglichkeit zur Weiterentwicklung von Produkten und Services bietet sich darüber hinaus durch die Verknüpfung mit Maschinellen Lernen und Künstlicher Intelligenz. Diese Technologien sind in der Lage, das digital vernetzte und globale Produkt zusätzlich noch mit einer weitgehenden Automatisierung und der Fähigkeit zum eigenständigen Lernen auszustatten. Die so entstehenden intelligenten Produkte und Services sind dadurch in der Lage, nicht nur komplexere Routineaufgaben selbständig durchzuführen, sondern auch Abweichungen vom Normalzustand selbständig zu erkennen und flexibel darauf zu reagieren.

Nachdem Disruptionen – also die Zerstörung eines alten Geschäftsmodells und Ersetzung durch ein überlegenes neues Geschäftsmodell – vor allem durch IT und dort insbesondere durch Maschinelles Lernen und Künstliche Intelligenz vorangetrieben werden, werden Smarte Produkte und die dazu gehörenden, durch IT-unterstützten, Dienstleistungen zahlreiche Geschäftsmodelle und ganze Märkte revolutionieren. Was genau Smarte Produkte und Dienstleistungen sind, welche Beispiele es dafür gibt und wie sie grundsätzlich funktionieren wurde in den beiden vorangegangenen Leitfäden des Mittelstand 4.0-Kompetenzzentrums Stuttgart beschrieben:

- ▶ **Smarte Produkte und Dienstleistungen**  
Wie sie als Entscheider Schritt für Schritt ins Thema einsteigen
- ▶ **IoT Architekturen für Entscheider\*innen**  
Make or Buy? – Eine Orientierungshilfe zum Einstieg in die digitale Transformation von Services und Produkten

An dieser Stelle möchten wir dagegen, die oben angesprochene Produkt-Service-Logik<sup>1</sup> erweitern, hin zu einer Produkt-Service-Digitalisierungs-Logik, die die Phasen der digitalen Produkt- und Serviceentwicklung veranschaulicht. Erstere unterscheidet zwischen dingbasierten und IT-basierten Innovationen. Dabei gehen aus den IT-basierten Innovationen vernetzte Produkte und Services hervor. Während dingbasierte Innovationen lediglich neue Produktfunktionen direkt und lokal am Produkt betreffen, führen IT-basierte Innovationen – durch Digitalisierung und Vernetzung – zu einer globalen Bereitstellung von Produkten und zugehörigen Services. Dies eröffnet nicht nur eine wesentlich größere Zielgruppe, sondern auch die Möglichkeit der Vernetzung mit anderen Services und der Aggregation von Funktionen. Somit werden wesentlich größere Potenziale für Kundenmehrwerte eröffnet als bei der dingbasierten Weiterentwicklung. Die Voraussetzung hierfür ist die Vernetzung der Ausgangsprodukte im Internet der Dinge (IoT).

In Abbildung 1 wird diese Logik ergänzt durch die KI-basierte Erweiterung. Das Produkt und die zugehörigen Services bleiben also in ihrer Entwicklung nicht stehen bei der Vernetzung und globalen Bereitstellung von Daten und Services. Stattdessen werden mit Hilfe der in den Produkten gewonnenen Daten und ggf. weiterer Daten aus anderen Bereichen KI-Systeme anhand von Verfahren des Maschinellen Lernens trainiert. Somit entstehen nicht nur global verfügbare vernetzte Produkte und Services, sondern darüber hinaus auch noch adaptive und autonom agierende Dienste. Der so gewonnene Mehrwert geht nochmal deutlich über die IT-basierten

## Produkt-Service-KI-Logik – neue Merkmale für Leistungsangebote

Ding + IT $\left\langle \begin{array}{l} \text{Hardware} \\ \text{Software} \end{array} \right.$		+ KI =	Ding-basierte Funktionen $\left\langle \begin{array}{l} \text{physisch} \\ \text{lokal} \end{array} \right.$	+ IT-basierte Services $\left\langle \begin{array}{l} \text{digital} \\ \text{global} \end{array} \right.$	+ KI-basierte Erweiterungen $\left\langle \begin{array}{l} \text{autonom} \\ \text{adaptiv} \end{array} \right.$
 Armbanduhr	 Sportuhr		<ul style="list-style-type: none"> <li>■ Zeit</li> </ul>	<ul style="list-style-type: none"> <li>■ Leistungsermittlung durch Messdatenerfassung und GPS-Streckenaufzeichnung</li> </ul>	<ul style="list-style-type: none"> <li>■ Leistungsdatenauswertung/Fitnesstests</li> <li>■ Trainingshinweise geben</li> </ul>
 Socken	 Sensoria-Fitnesssocken		<ul style="list-style-type: none"> <li>■ Wärme</li> <li>■ Laufkomfort</li> </ul>	<ul style="list-style-type: none"> <li>■ Entwicklungsplattform für textile Elektronik</li> <li>■ Bluetooth-Anbindung</li> <li>■ Sensordatenaufnahme</li> </ul>	<ul style="list-style-type: none"> <li>■ Verbesserung schlechter Laufgewohnheiten durch Auswertung der Gewichtsverteilung und deren Auswirkungen</li> </ul>
 Drehbank	 Drehautomat		<ul style="list-style-type: none"> <li>■ Produktionsverfahren Drehen</li> </ul>	<ul style="list-style-type: none"> <li>■ Optimierung der Produktionsplanung durch vernetzte Maschinen und Anlagen</li> </ul>	<ul style="list-style-type: none"> <li>■ Optimierung Wartungsprozesse durch Predictive Maintenance</li> <li>■ Bedienhinweise durch Sensordatenauswertung</li> </ul>
Versicherungsmakler	Robo-Advisor		<ul style="list-style-type: none"> <li>■ Klassischer Versicherungsvertrieb</li> </ul>	<ul style="list-style-type: none"> <li>■ Digitaler Versicherungsvertrieb über die Websites von Versicherung (digitaler Vertragsassistent)</li> </ul>	<ul style="list-style-type: none"> <li>■ KI-basierter Vergleich von Versicherungsnehmern anhand ihrer Profile</li> </ul>

Abbildung 1: Produkt-Service-KI-Logik

Die in Abbildung 1 dargestellte Klassifizierung nennen wir Produkt-Service-KI-Logik. Beispiele für entsprechende Erweiterungen können selbstlernende Produkte sein, ihre Konfiguration an den aktuellen Anwendungsfall oder den Anwender durch Analyse vergangener Anwendungsfall-, Anwender- und Ergebnis-Kombinationen selbständig anpassen. Damit gehen sie über die sensorgestützte Erfassung und Auswertung von Nutzungsdaten oder Benutzer-Hinweise zur Verbesserung der Produkthandhabung hinaus. Sie erlauben die Personalisierung von Produkten und die Differenzierung gegenüber dem Wettbewerb. KI-basierte Personalisierungsansätze eröffnen die Möglichkeit zu einer neuen Form des Lernens und zur gezielten individuellen Anpassung von Produkten und Dienstleistungen. Leistungsangebote und Service-Qualität können auf diese Weise auf eine neue Ebene gehoben werden und zielgerichtet die jeweils individuell benötigten Mehrwerte liefern. Ein Beispiel ist die

Entwicklung eines KI-basierten digitalen Assistenzsystems<sup>2</sup> für Studierende, wie es von der Fern-Uni Hagen in Zusammenarbeit mit dem DFKI entwickelt wird. Studierende sollen damit eine Art persönlichen Begleiter erhalten, der ihnen durchs Studium hilft und beispielsweise auch Empfehlungen für individuelle Lernstrategien gibt. Ein weiteres Beispiel ist die Anwendung von Text-Mining-Techniken in der Marktforschung<sup>3</sup>, um Kundeninformationen aus großen Datenmengen (z. B. Kundenfeedback oder Berichte von Servicetechnikern) zu extrahieren sowie Kundenerfahrungen zu messen und auszuwerten. Durch das automatische Extrahieren und Strukturieren von Informationen können Unternehmen das Kundenerlebnis auf intelligente Weise bewerten und das Leistungsangebot darauf aufbauend durch Vereinfachung und Automatisierung von Geschäftsprozessen verbessern. Dies erfolgt z. B. beim digital unterstützten Reklamationsmanagement oder im Kontext der Vorausschauenden Wartung von Produkten (engl. Predictive Maintenance). Auch Serviceanfragen können mittels KI besser bearbeitet werden, indem die KI die Anfragen automatisiert bestimmten Servicekategorien zuordnet und dann an entsprechende Sachbearbeiter weiterleitet<sup>4</sup>.

## 4. Wie funktioniert künstliche Intelligenz?

Mit dem Begriff »Künstliche Intelligenz« werden Computersysteme bezeichnet, die zumindest teilweise menschenähnliche Intelligenzleistungen vollbringen. Doch was genau ist Intelligenz und welche Arten von Intelligenzleistungen sind hiermit gemeint? Das Problem bei allen Definitionen künstlicher Intelligenz ist der – trotz vielfältiger Versuche – nach wie vor nicht ausreichend geklärte Begriff der Intelligenz. Was genau ist es, das Menschen – oder eben ggf. auch Computersysteme oder Maschinen – intelligent macht?

Mit dieser Frage hat sich eine Expertenkommission der Europäischen Kommission befasst und ist zu folgender [Definition von künstlicher Intelligenz](#) gekommen:

„Systeme der künstlichen Intelligenz (KI-Systeme) sind vom Menschen entwickelte Softwaresysteme (und gegebenenfalls auch Hardwaresysteme), die in Bezug auf ein komplexes Ziel auf physischer oder digitaler Ebene handeln, indem sie ihre Umgebung durch Datenerfassung wahrnehmen, die gesammelten strukturierten oder unstrukturierten Daten interpretieren, Schlussfolgerungen daraus ziehen oder die aus diesen Daten abgeleiteten Informationen verarbeiten, und über das bestmögliche Handeln zur Erreichung des vorgegebenen Ziels entscheiden. KI-Systeme können entweder symbolische Regeln verwenden oder ein numerisches Modell erlernen, und sind auch in der Lage, die Auswirkungen ihrer früheren Handlungen auf die Umgebung zu analysieren und ihr Verhalten entsprechend anzupassen.

Als wissenschaftliche Disziplin umfasst die KI mehrere Ansätze und Techniken wie z. B. maschinelles Lernen (Beispiele dafür sind „Deep Learning“ und bestärkendes Lernen), maschinelles Denken (es umfasst Planung, Terminierung, Wissensrepräsentation und Schlussfolgerung, Suche und Optimierung) und die Robotik (sie umfasst Steuerung, Wahrnehmung, Sensoren und Aktoren sowie die Einbeziehung aller anderen Techniken in cyber-physische Systeme).“

Eine weitere Herangehensweise an das Thema künstliche Intelligenz ist die Auseinandersetzung mit den Anwendungsgebieten. Künstliche Intelligenz wird somit definiert über die Dinge, die sie zu leisten in der Lage ist. Das vom Deutschen Digitalverband BITKOM unterstützte [Periodensystem KI](#) bietet eine Übersicht über die Anwendungsgebiete, beschreibt was genau sich dahinter verbirgt und auch wie die einzelnen Anwendungsgebiete zusammen- und voneinander abhängen. Der Vorteil dieser Betrachtungsweise liegt für Nutzer von KI-basierten Dienstleistungen unter anderem darin, dass man die verschiedenen KI-Angebote besser einordnen kann. Es ist beispielsweise möglich zu erkennen, ob ein Dienstleister nur einen Teil meines Problems als potenzieller Kunde oder den gesamten Lösungsraum mit abdecken kann, indem man die Zusammenhänge zwischen Teilgebieten und entsprechenden Angeboten besser einordnen kann.

Die wichtigsten Merkmale von KI-Systemen sind Autonomie (Selbständigkeit) und Adaptivität (Anpassungsfähigkeit). KI-Anwendungen sind also in der Lage, gewisse – im [Periodensystem KI](#) zusammengefasste – Aufgaben ohne menschliche Hilfe zu erledigen und sich ggf. an unvorhergesehene Umgebungsbedingungen anzupassen – obwohl ihnen vorher niemand konkret beigebracht hat, wie sie sich in genau dieser Situation verhalten sollten.

Ein weiteres Nachschlagewerk für ein besseres Verständnis der künstlichen Intelligenz ist der [Praxisleitfaden Künstliche Intelligenz für den Mittelstand](#) des Mittelstand 4.0-Kompetenzzentrums Kaiserslautern. Dort werden vor allem auch Praxisbeispiele aus kleinen und mittleren Unternehmen dargestellt und verständlich erklärt.

Die künstliche Intelligenz wird im Wesentlichen in drei Teilbereiche untergliedert:

- ▶ Künstliche Intelligenz (Artificial Intelligence)
- ▶ Maschinelles Lernen (Machine Learning)
- ▶ Tiefes Lernen (Deep Learning)

Die Zusammenhänge zwischen den drei Gebieten sind in Abbildung 2 dargestellt.



Abbildung 2: Unterarten Künstlicher Intelligenz

Von den 1950er Jahren bis in die 1980er Jahre standen nur regelbasierte Möglichkeiten zur Verfügung. Mit dem maschinellen Lernen konnten Computer dann ab den 1980er Jahren anfangen, aus vorhandenen Daten zu lernen, ähnlich wie Menschen das tun. Mit den Möglichkeiten des Tiefen Lernens (engl.: Deep Learning) konnten in den letzten 10 Jahren erheblich komplexere Lernvorgänge umgesetzt werden, was der künstlichen Intelligenz einen deutlichen Entwicklungsschub gegeben hat. Ein Grund für diesen Entwicklungssprung ist auch die Entwicklung der Computer-Hardware, die die Berechnung von hochkomplexen neuronalen Netzen erst ermöglicht hat. Neben der Hardware und den Algorithmen sind für die Umsetzung von KI-Systemen vor allem geeignete Daten für das Training der Algorithmen und entsprechende Entwickler erforderlich. Data Scientists müssen

nicht nur in der Lage sein, den Maschinen das Lernen beizubringen, sondern auch fähig sein, Daten zu verwalten, zu analysieren, aufzubereiten und für eine hohe Datenqualität in den Datensätzen zu sorgen. Hierbei spielen auch organisatorische Voraussetzungen in den Unternehmen eine Rolle. Es ist also nicht damit getan, leistungsfähige Hardware mit leistungsfähigen Algorithmen einzukaufen.

Hinzu kommt, dass je nach Tiefe des Maschinellen Lernens auch immer größere Mengen an Daten für das Lernen benötigt werden. Da die Datenmengen sich exponentiell entwickeln ist diese Voraussetzung in immer mehr Fällen gegeben. Auch dies hat dazu beigetragen, dass sich das Tiefe Lernen (engl.: Deep Learning) im letzten Jahrzehnt rasant weiterentwickelt hat.

## 5. Maschinelles Lernen - wie geht das?

Im Prinzip funktioniert maschinelles Lernen so ähnlich wie das menschliche Lernen als Kind. Wenn man geboren wird weiß man nicht was Hunde und Katzen sind. Aber mit jedem Hund und jeder Katze, die man zu Gesicht bekommt – und dem Hinweis der Eltern, dass das jetzt ein Hund oder eine Katze war – lernt man Hunde und Katzen voneinander zu unterscheiden. Ähnlich funktioniert das auch mit Computern:

Man gibt dem Computer Daten, aus denen er lernen kann. Ähnlich wie unser Gehirn Synapsen knüpft, entwickelt der Computer dann selbständig Algorithmen, um das Gelernte zu beschreiben.

### Supervised vs. Unsupervised

Man unterscheidet dabei Überwachtes und Unüberwachtes Lernen (engl.: supervised and unsupervised learning). Am Beispiel der Hunde und Katzen bedeutet überwachtes Lernen, dass die Eltern dem Kind (oder die Programmierer dem Computer) sagen, auf welchem Bild ein Hund und auf welchem Bild eine Katze zu sehen ist. Die Alternative dazu wäre, dass man selbst auf den Unterschied kommen müsste und aus den Ähnlichkeiten der Hunde und der Katzen selbst zwei Gruppen bildet. Natürlich funktioniert es zuverlässiger wenn einem jemand die richtige Lösung sagt – oder in der Sprache des Maschinellen Lernens ausgedrückt: wenn der Datensatz die Zielvariable enthält.

### Anwendungsbeispiele

Maschinelles Lernen kann wie im Beispiel oben zur Unterscheidung von Katzen und Hunden in Bildern eingesetzt werden oder – in einem industriellen Kontext – zur Unterscheidung von intakten Produkten und beschädigten Produkten im Zuge der Qualitätssicherung in der Produktion. Es können auch individuelle Produktvorschläge im Online-Shop sein, bei denen die KI im Online-Shop aus meinem Nutzungsverhalten der Vergangenheit und meiner Ähnlichkeit zu anderen Nutzern (...der verhält sich ja genau wie Max Mustermann...) gelernt hat vorherzusagen, was mir sonst noch gefallen könnte – nämlich dasselbe wie Max Mustermann. Es kann auch die Vorausschauende Wartung sein, bei der man so lange Produkte und Produkteigenschaften beobachtet, bis man vorhersagen kann, wann sie voraussichtlich kaputtgehen. Die Hinweise können beispielsweise Veränderungen in der Geräuschkulisse sein, die auf Reibungen oder Vibrationen hindeuten und diese wiederum weisen dann auf ein baldiges Versagen des Produkts hin. Auf diese Weise kann dann beispielsweise der Aufzug repariert werden, bevor jemand darin steckenbleibt.

## Datengrundlage

Die Voraussetzung für Maschinelles Lernen ist immer, dass man genügend Daten hat, aus denen man lernen kann und dass diese Daten von ausreichend hoher Datenqualität sind. Dann muss man Parameter finden, die eine Unterscheidung erlauben zwischen den Zuständen, die man unterscheiden möchte (Hund oder Katze, intakt oder kaputt, gesund oder krank, attraktiv oder nicht, etc.). Wenn die Daten oder Kombinationen der Daten es erlauben, den Unterschied zu erkennen indem ein Zusammenhang zwischen den Daten und dem zu unterscheidenden oder vorherzusagenden Merkmal (Zielvariable) besteht, wird auch der Computer lernen können, diesen Unterschied zu erkennen. Die Algorithmen, mit denen er den Unterschied dann erkennt, müssen nicht mehr von Hand von einem Menschen entwickelt werden. Sie entwickeln sich während des Lernens selbst. Unterstützt werden die KI-Entwickler von verschiedenen Lern-Modellen, die je nach Aufgabe mal besser und mal schlechter geeignet sind. Am Ende kommt es darauf an, mit welcher Genauigkeit die KI Zustände voneinander unterscheiden kann.

Abbildung 3 gibt eine Übersicht über die wichtigsten Modelle des Maschinellen Lernens, ohne diese näher erläutern zu wollen. Die Begriffsübersicht soll hier als Ausgangspunkt für die selbständige Vertiefung dienen.

## Übertragbarkeit

Wenn das Modell gelernt hat, gut von schlecht und falsch von richtig zu unterscheiden, hat es eine Entscheidungsgrundlage um auch in unvorhergesehenen Situationen die richtige Wahl zu treffen und sich selbständig (Autonomie) an so eine Situation anzupassen (Adaptivität). In einem klassischen regelbasierten System müsste man dem Computer Bilder aller möglichen Hunderassen und Hunde auf der Welt geben, damit es Hunde sicher identifizieren könnte. Durch Maschinelles Lernen erübrigt sich dieser Aufwand und der Computer findet mit der Zeit selbst heraus, woran Hunde zu erkennen sind – ähnlich wie das auch jedes Kind irgendwann selbst herausfindet. Und wenn ein Computer das einmal herausgefunden hat, dann kann man diese Erkenntnis auf alle Computer übertragen, indem man den Algorithmus dorthin kopiert. Während jeder Mensch für sich lernen muss wie Hunde und Katzen zu unterscheiden sind, da das Kopieren zwischen menschlichen Gehirnen nicht möglich ist, reicht es im Prinzip aus, wenn ein Computer etwas einmal gelernt hat. Anschließend steht der Übertragung auf alle Computer nichts mehr im Weg – sofern der KI-Entwickler bereit ist, den Algorithmus mit anderen zu teilen...

## Grundsätzlicher Ablauf

Die prinzipielle Vorgehensweise beim Maschinellen Lernen ist immer die folgende:

Zunächst werden Daten gesammelt – z.B. über die Sensorik in Produkten – anhand derer man etwas lernen kann. Anschließend muss man die Daten aufbereiten und beispielsweise unvollständige oder fehlerhafte Daten aussortieren oder dafür sorgen, dass alle Daten im selben Maßstab vorliegen. Es soll beispielsweise schon Raumsonden gegeben haben, die auf fremden Planeten abgestürzt sind, weil Meter mit Inch verwechselt wurden...

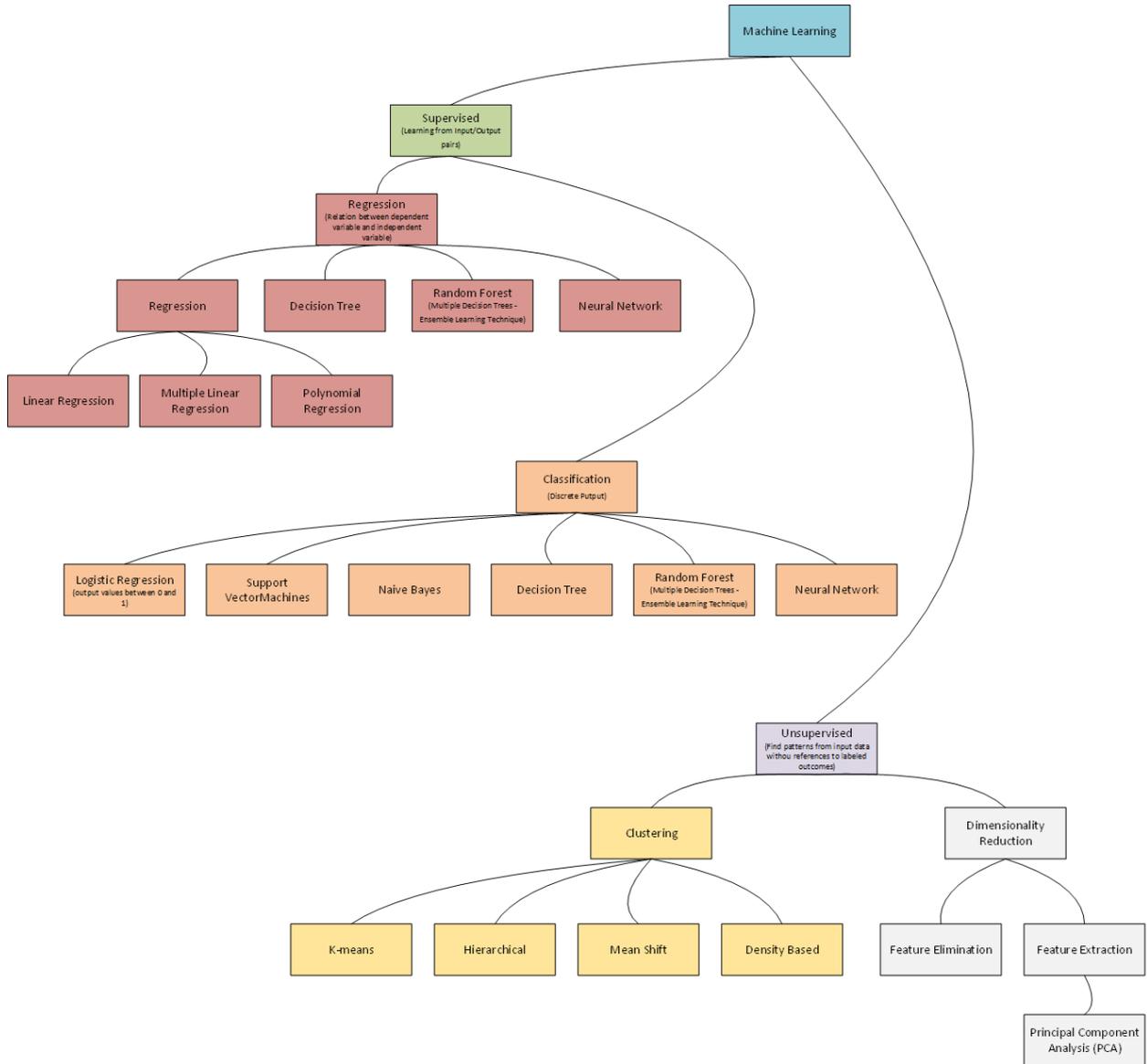


Abbildung 3: Übersicht über die wichtigsten Arten von Modellen des Maschinellen Lernens

Wenn die Daten dann in guter Qualität und einem sinnvollen Format vorliegen wird ein Datenmodell erstellt. Mit Hilfe dieses Datenmodells wird dann ein (bzw. mehrere) Modell(e) des Maschinellen Lernens (siehe Abbildung 3) trainiert und anschließend überprüft bzw. getestet. Um die Überprüfung durchführen zu können, benötigt man einen anderen Datensatz als den, mit dem man das ML-Modell trainiert hat. Ansonsten kann es passieren, dass die Maschine zwar lernt alle Hunde und Katzen in dem Trainings-Datensatz zu unterscheiden, aber an anderen Hunden und Katzen in der realen Welt scheitert. Um dies zu vermeiden, trennt man seinen Ausgangsdatsatz in Trainingsdaten und Testdaten. Dabei sollten die beiden Datensätze eine vergleichbare Wahrscheinlichkeitsverteilung haben. Das heißt, dass die Unterscheidung des gewünschten Merkmals in beiden Sätzen gleich gut funktionieren sollte und nicht in einem davon besser oder schlechter.

Abschließend wird dann das ML-Modell anhand des Testdatensatzes geprüft. In der Praxis werden dabei meist viele Iterationsschleifen durchlaufen bis das Modell schließlich eine ausreichend hohe Genauigkeit aufweist. Wenn die KI in einem von Hundert Fällen einen Fehler machen darf, muss man weniger lang trainieren als wenn sie nur in einem von Tausend Fällen einen Fehler machen darf.

In dem Industrieszenario mit der Unterscheidung zwischen qualitativ guten Produkten und auszusortierenden Produkten entscheidet die Genauigkeit über bares Geld, da mit jedem ausgelieferten Produkt, das beanstandet wird, ein aufwendiger Kundenservice-Prozess und der Austausch des Produkts verbunden ist. Ein Extrembeispiel ist hier die extrem teure Rückrufaktion in der Automobilindustrie, weil z. B. schlecht produzierte Bremsen in der Qualitätssicherung nicht zuverlässig aussortiert werden konnten. Personenschäden aufgrund fehlerhafter Bremsen oder der Schaden für den Ruf des Herstellers sind dabei noch gar nicht berücksichtigt.

### Kontinuierliches Lernen

Eine weitere Unterscheidung der Ansätze für das maschinelle Lernen ist, ob man eine KI-Anwendung auf einem festen Datensatz einmal trainiert, wie das beispielsweise für die Unterscheidung von Hunden und Katzen durchaus Sinn ergibt, oder ob man das Modell kontinuierlich mit immer neuen Daten weiter trainiert. Letzteres ergibt z.B. in dem Szenario mit dem Online-Shop und den Produktempfehlungen Sinn. Wenn ich Max Mustermann ähnele und der ein neues Produkt kauft, das zur Zeit des Trainings der KI noch gar nicht auf dem Markt

war, kann mir die KI das Produkt gar nicht vorschlagen. Wenn sie aber permanent dazu lernt, weil das Modell mit immer neuen Daten (also dem Kaufverhalten von Max Mustermann) wieder und wieder trainiert wird, kann sie mir auch die neueste Errungenschaft von Herrn Mustermann vorschlagen. Und im Idealfall (aus Sicht des Online-Händlers) lege ich das Produkt aufgrund des Vorschlags dann in meinen Warenkorb und kaufe es. Und wenn der Zusammenhang zwischen den Produktvorschlägen und meinem Kaufverhalten ausreichend gut ist rentiert sich auch die Investition in das Maschinelle Lernen früher oder später. Das Training des Modells muss also – wenn es einmal grundsätzlich vernünftige Ergebnisse liefert – automatisiert mit aktualisierten Daten wiederholt werden und die eigentliche KI-Anwendung muss laufend aktualisiert werden.

## 6. IT-Architekturen für Maschinelles Lernen

Im Zuge der Umsetzung intelligenter Produkte und Dienstleistungen spielen Daten und damit auch die Datenverarbeitung eine zentrale Rolle. Die wesentlichen Elemente der Datenverarbeitung bei der Nutzung Künstlicher Intelligenz bzw. beim Maschinellen Lernen sind in Abbildung 4 dargestellt. Der Datenverarbeitungsprozess startet unten im Bild bei der Anbindung verschiedenster Datenquellen über geeignete Schnittstellen. Die Daten werden kontinuierlich in die KI-IT-Architektur eingespeist, z.B. indem Smartphones ihre Positionsdaten in gewissen zeitlichen Intervallen an eine bestimmte Adresse schicken und das Gerätemanagement in der Cloud, die sich hinter dieser Adresse verbirgt, die Daten in einen entsprechenden Datenverarbeitungsprozess einspeist. Es entsteht also ein Datenfluss. Dieser wird angereichert mit weiteren Informationen aus eigenen oder auch fremden IT-Anwendungen, z.B. mit Geoinformationen aus einem Kartendienst. Der Datenfluss wird in einem Datenspeicher verwaltet. Dieser kann ein klassisches Data Warehouse oder alternativ – im Falle sehr großer Datenmengen – auch ein Data Lake<sup>5</sup> sein. Es folgen dann diverse Schritte der Datenvorverarbeitung und Datenaufbereitung (Data Pre-Processing). So müssen Daten beispielsweise in geeignete Formate umgewandelt werden und auch auf ein Datenqualitätsmanagement ist an dieser Stelle zu achten, um die KI später nur mit validen und sinnvollen Daten zu füttern. Ist die Vorverarbeitung abgeschlossen, müssen die Daten in verschiedene Datensätze aufgeteilt werden. Um einer KI das Lernen zu ermöglichen werden am besten drei verschiedene Datensätze benötigt, nämlich für das Training der KI, für die Validierung und Überprüfung der erlernten KI-Modelle und für ausführliche Tests. Dabei ist darauf zu achten, dass keine qualitativen Unterschiede zwischen den verschiedenen Datensätzen entstehen, die beispielsweise dazu führen könnten, dass die wichtigen Informationen nur im Trainingsdatensatz enthalten sind und die unwichtigen alle im Validierungsdatensatz. Ein anschauliches Beispiel wie es nicht laufen sollte ist, um bei dem Beispiel mit den Hunde- und Katzenbildern zu bleiben, wenn man einen Gesamtdatensatz mit Dackeln, Pudeln und Schäferhunden hat und ihn anschließend in einen Trainingsdatensatz mit Dackelbildern, einen Validierungsdatensatz mit Pudelbildern und einen Testdatensatz mit Bildern von Schäferhunden aufteilt. Um Hunde und Katzen zuverlässig unterscheiden zu können, sollte die richtige Verteilung von Hunderassen in allen drei Datensätzen vorkommen. Dasselbe gilt natürlich auch für die verschiedenen Arten von Katzen.

## IoT Architektur für Machine Learning

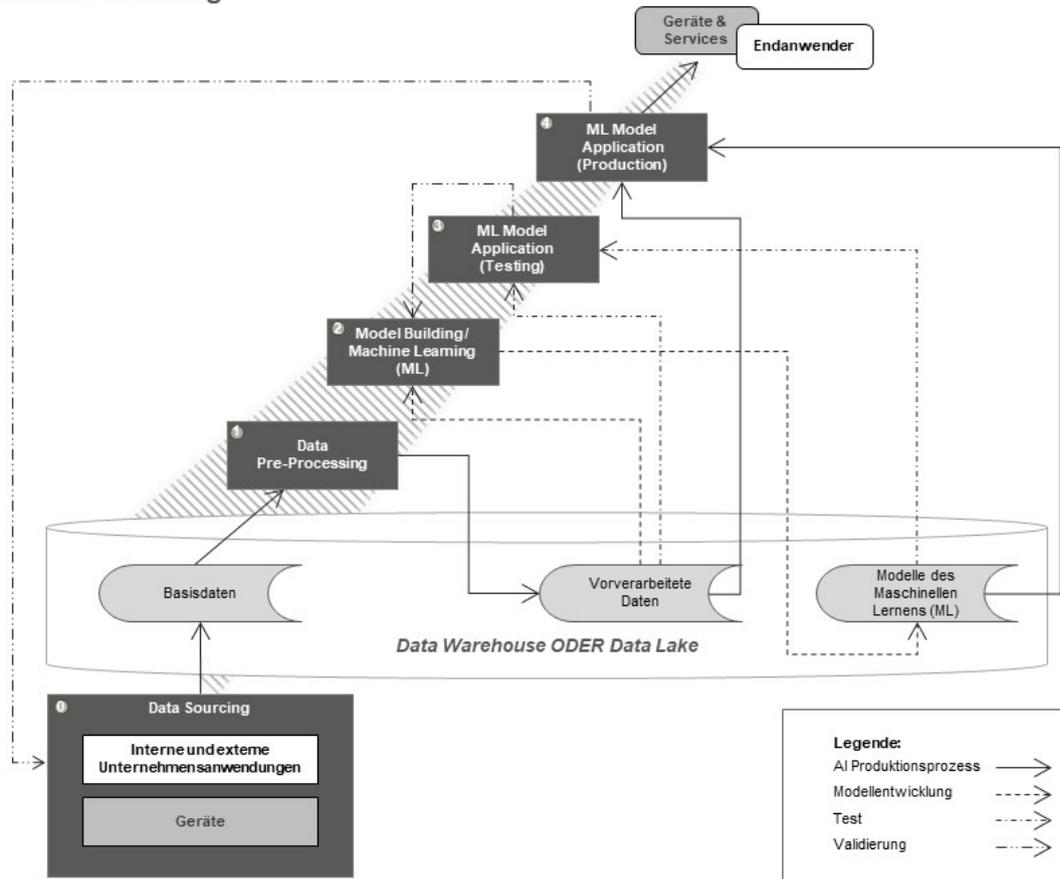


Abbildung 4: Wesentliche Bestandteile einer Machine Learning-orientierten IT-Architektur

Nach der Aufteilung der Datensätze beginnt der Prozess der sogenannten Modellbildung, zu der auch das Maschinelle Lernen gehört. Hier werden zunächst aus den Trainingsdaten ML-Modelle abgeleitet, die dann später von der KI-Anwendung genutzt werden, um in Zukunft aus Livedaten von Geräten oder aus Kundenprozessen selbständig und selbstlernend sinnvolle Lösungen anbieten zu können. Die ersten Modell-Entwürfe werden dann mit den Validierungsdatensätzen überprüft, um das Modell zu verfeinern. Wenn das Modell dann die gewünschte Genauigkeit liefert – beispielsweise beim Erkennen bestimmter Bilder – wird das Modell in einer Test-Anwendung implementiert und mit dem Test-Datensatz nochmal ausführlich überprüft. Mit den Rückmeldungen aus den Testläufen können die Modelle nochmals schrittweise verfeinert werden. In einem letzten Schritt werden die Modelle dann in den finalen Anwendungen implementiert. Diese werden dann mit den Geräte- und Prozessdaten der Kunden gespeist, um die eigentliche KI-Anwendung zu erbringen.

Um den Gesamtaufwand einer entsprechenden Umsetzung abschätzen zu können sollte man sich aber vergegenwärtigen, dass die in Abbildung 4 dargestellten Schritte sich ausschließlich auf die Kernprozesse der KI-Umsetzung beziehen. Diese sind eingebettet in komplexere IT-Architekturen, wie sie für Produkte und Dienstleistungen im Internet der Dinge erforderlich sind.

Abbildung 5 zeigt eine Übersicht über die wichtigsten Schichten einer IoT-Architektur<sup>6</sup>.

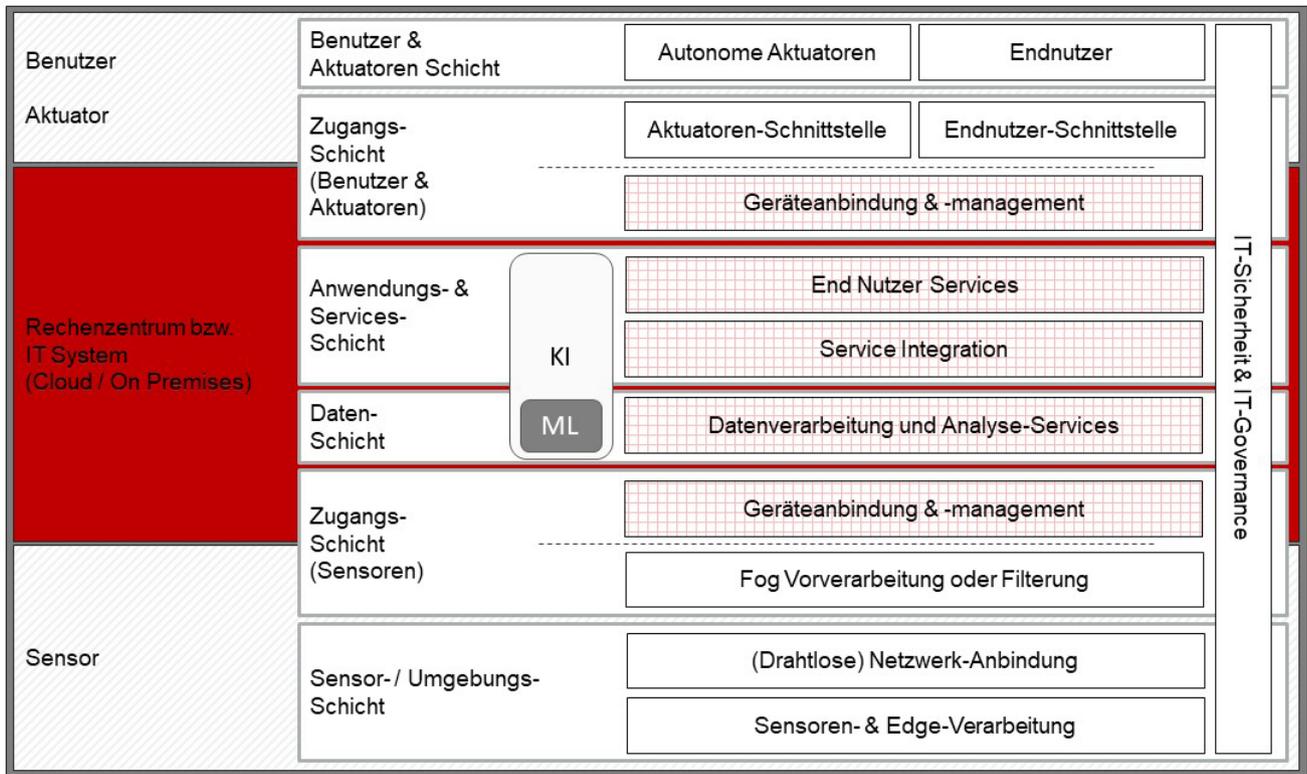


Abbildung 5: Schichten einer IoT-Architektur

## Sensoren und Edge-Verarbeitung

Ganz unten im Bild ist dabei die Ebene der Sensorik zu finden, die in die Produkte verbaut ist und in der Regel den Ausgangspunkt der Datenerfassung bildet. Auf dieser Ebene findet auch ggf. bereits eine Vorverarbeitung von Daten statt, beispielsweise um Datenmengen zu reduzieren bevor sie in die weiterverarbeitenden IT-Systeme geschickt werden. Direkt darüber befindet sich die Schicht der Netzwerkanbindung. Hier findet die Vernetzung mit den weiterverarbeitenden IT-Systemen statt, beispielsweise über WLAN oder Mobilfunktechnologien. Noch im lokalen Netz ist optional eine weitere Vorverarbeitung möglich, alternativ zur Vorverarbeitung direkt auf

Beide genannten Vorverarbeitungsschritte in Abbildung 5 werden durchgeführt, um entweder Datenmengen zu reduzieren, weil die Datenanbindung an die Cloud oder ein eigenes Rechenzentrum (RZ) keine ausreichende Bandbreite bietet, oder weil Echtzeitanforderungen dazu führen, dass die Signallaufzeiten in die Cloud / ins Rechenzentrum und wieder zurück zu nicht tolerierbaren Verarbeitungszeiten führen. Man unterscheidet hierbei die Vorverarbeitung direkt auf dem Gerät bzw. Edge Computing, und die Vorverarbeitung auf lokalen IT-Komponenten außerhalb der Geräte, das sogenannte Fog Computing<sup>7</sup>.

## IT-Systeme, Rechenzentren, Clouds

Anschließend verlassen die Daten das lokale Umfeld der Geräte und werden in das weiterverarbeitende IT-System geschickt. Hierfür bietet sich eine skalierbare Cloud-Lösung an, um auch mit steigenden Kunden- und Gerätezahlen immer ausreichend Leistung zur Verfügung zu haben, ohne ständig die IT nachjustieren zu müssen. In diesem mittleren Block beginnt der Prozess mit der Geräteanbindung (Device Connection & Management), deren Skalierbarkeit insbesondere bei größeren Geräteanzahlen entscheidend ist und die auch den Hauptansatzpunkt liefert um umgekehrt wieder Geräte steuern zu können – beispielsweise durch Over the Air (OTA) Updates. Anschließend folgt dann die eigentliche Datenverarbeitung (Data Processing & Analytics Services). In diesem Block wären auch die Kernkomponenten der oben beschriebenen KI einzuordnen, die an der Stelle beginnen, wo die Geräte über die Geräteanbindung ihre Daten beispielsweise in einen Data Lake in der Cloud einspeisen. Auf die – in unserem Fall auf Maschinellem Lernen basierende – Datenanalyse folgt die Integration der verschiedenen Services, die auf die verarbeiteten Daten zugreifen, bevor diese End-Nutzer-Services

dann schließlich die eigentlichen Anwendungen für Kunden, Partner, den Anbieter selbst oder für Aktorik in den Geräten zur Verfügung stellen.

### Benutzer und Aktuatoren

Nachdem die Geräteanbindung ggf. ein weiteres Mal – nur diesmal in umgekehrter Richtung – durchlaufen wird, um zu senden statt zu empfangen, gelangen die für das Funktionieren der Produkte oder die Erbringung von Dienstleistungen erforderlichen Ergebnisse dann an die bereits genannten Endverbraucher: Kunden, Partner, Anbieter oder Aktorik in den Geräten.

### Cloudbasierte KI-Architekturen

Einen weiterführenden Einblick in die Möglichkeiten cloudbasierter KI-Architekturen ermöglicht die [Studie Cloudbasierte KI-Plattformen](#), die von Fraunhofer im Rahmen des KI Fortschrittszentrums Lernende Systeme in 2021 veröffentlicht wurde. Sie bietet vor allem für Unternehmen, die sich keine eigenen IoT- oder KI-Infrastrukturen leisten können, eine Alternative um eigene KI-Services entwickeln zu können.

## Ihr Kontakt zu uns

- ▶ Ausführliche Informationen zum Mittelstand 4.0-Kompetenzzentrum Stuttgart, einen Einblick in unser gesamtes Themenspektrum und aktuelle Terminhinweise finden Sie unter: [digitales-kompetenzzentrum-stuttgart.de](https://digitales-kompetenzzentrum-stuttgart.de).
- ▶ Sie haben eine konkrete Anfrage zum Themenfeld Cloud Computing und Smart Services? Senden Sie eine Mail an unseren Verantwortlichen. Er wird Ihre Anfrage an die entsprechenden Ansprechpartner\*innen weiterleiten, die sich dann mit Ihnen in Verbindung setzen werden.



Jürgen Falkner

Verantwortlicher für das Thema „Cloud Computing  
und Smart Services“

[cloud@digitales-kompetenzzentrum-stuttgart.de](mailto:cloud@digitales-kompetenzzentrum-stuttgart.de)

Wir freuen uns auf Ihre Anfrage und eine erfolgreiche Zusammenarbeit  
Ihr Team vom Mittelstand 4.0-Kompetenzzentrum Stuttgart

## Was ist Mittelstand-Digital?

Mittelstand-Digital informiert kleine und mittlere Unternehmen über die Chancen und Herausforderungen der Digitalisierung. Die geförderten Kompetenzzentren helfen mit Expertenwissen, Demonstrationszentren, Best-Practice-Beispielen sowie Netzwerken, die dem Erfahrungsaustausch dienen. Das Bundesministerium für Wirtschaft und Energie (BMWi) ermöglicht die kostenfreie Nutzung aller Angebote von Mittelstand-Digital. Der DLR Projektträger begleitet im Auftrag des BMWi die Kompetenzzentren fachlich und sorgt für eine bedarfs- und mittelstandsgerechte Umsetzung der Angebote. Das Wissenschaftliche Institut für Infrastruktur und Kommunikationsdienste (WIK) unterstützt mit wissenschaftlicher Begleitung, Vernetzung und Öffentlichkeitsarbeit. Weitere Informationen finden Sie unter [www.mittelstand-digital.de](http://www.mittelstand-digital.de).

# Quellenverzeichnis

- 1 [Fleisch et al.: Geschäftsmodelle im Internet der Dinge](#)
- 2 [AI.EDU Research Lab](#)
- 3 [McColl-Kennedy et al.: Gaining Customer Experience Insights That Matter](#)
- 4 [Effizienter Kundendienst durch Kategorisierung von Serviceanfragen](#)
- 5 [Madera/Laurent: The Next Information Architecture Evolution: The Data Lake Wave](#)
- 6 [Kutzias et al.: On the Complexity of Cloud and IoT Integration: Architectures, Challenges and Solution Approaches](#)
- 7 [Yousefpour et al.: All one needs to know about fog computing and related edge computing paradigms: A complete survey](#)

## Bildnachweis

Titelseite:

# 220790568 © fotomek - AdobeStock

Seite : 8

© Erweiterung basierend auf (Fleisch et al. 2012)

Seite : 12

© Jürgen Falkner Fraunhofer IAO

Seite : 17

© Jürgen Falkner Fraunhofer IAO

Seite : 21

© Jürgen Falkner Fraunhofer IAO

Seite : 23

© Jürgen Falkner Fraunhofer IAO

Zur übersichtlicheren Darstellung wurde nur die maskuline Form verwendet. Selbstverständlich sind damit alle Geschlechter gemeint.

## Impressum

### Herausgeber

Mittelstand 4.0-Kompetenzzentrum Stuttgart  
c/o Fraunhofer-Institut für Arbeitswirtschaft und Organisation IAO  
Nobelstraße 12  
70569 Stuttgart

### Autoren

Jürgen Falkner, Fraunhofer IAO

### Rechtsform

Das Fraunhofer-Institut für Arbeitswirtschaft und Organisation IAO ist eine rechtlich nicht selbständige Einrichtung der Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e.V.  
Hansastraße 27 c  
80686 München  
Telefon: +49 89 1205-0  
Fax: +49 89 1205-7531  
[www.fraunhofer.de](http://www.fraunhofer.de)

### Druck

Fraunhofer Verlag  
Mediendienstleistungen  
Nobelstraße 12  
70569 Stuttgart

### Stand

Juni 2022

