

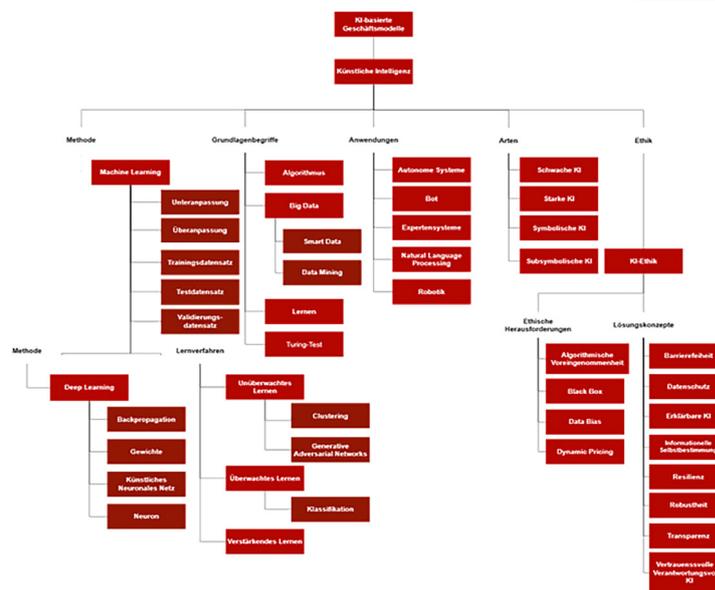
# Begriffe rund um Künstliche Intelligenz

## Faktenblatt

### Im Fokus

Der Begriff „Künstliche Intelligenz (KI)“ ist allgegenwärtig. KI ist ein Teilgebiet der Informatik und beschreibt die Fähigkeit einer Maschine, menschenähnliche Lern- und Denkmuster zu erbringen, indem sie Informationen auf Basis von Eingabedaten erkennt und interpretiert.

In der nachfolgenden Übersicht sind die wichtigsten Begriffe und Themen rund um diese Schlüsseltechnologie kurz und knapp erläutert.



## Algorithmus

Algorithmen sind (komplexe) Regelwerke zur Lösung von Problemen. Sie verarbeiten Eingabedaten mit Hilfe von verschiedenen mathematischen und statistischen Methoden, um auf diesem Weg eine Ausgabe in Form von Daten oder Handlungsvorschriften zu generieren.

## Algorithmische Voreingenommenheit

Algorithmische Voreingenommenheit (Algorithmic Bias) beschreibt, wie Vorurteile und strukturelle Diskriminierungen aus der realen

Welt in KI-basierte Systeme implementiert werden.

Bei dieser Art von Bias werden die Vorurteile in der Regel explizit durch die Entwickler:innen im Rahmen der Programmierung des jeweiligen Algorithmus hinterlegt. Auf diese Weise können sich bestimmte Merkmale negativ auf die Entscheidung bzw. Beurteilung durch das KI-System auswirken, auch wenn die zu Grunde liegenden Daten keine Vorurteile enthalten.

Formen von Diskriminierung, welche dem System implizit durch vorurteilsbehaftete Daten angeleitet werden, können unter dem Begriff Data Bias zusammengefasst werden.

## Autonome Systeme

Autonome Systeme können komplexe Aufgaben unter verschiedenen Ausgangssituationen lösen, indem sie selbstständig – ohne Hilfe von außen – einen Handlungsplan zur Problemlösung generieren und diesen umsetzen. Scheitert ein Teilabschnitt des Handlungsplans, ist das autonome System durch Methoden der künstlichen Intelligenz in der Lage, den Vorfall zu analysieren und Schlussfolgerungen zu ziehen, die in Zukunft zu einer Verhaltensanpassung des Systems führen können.

## Backpropagation

Der Backpropagation-Algorithmus stellt eine der zentralen Methoden für das Training neuronaler Netze dar. Dabei wird das vom KI-System ausgegebene Ergebnis, welches zunächst auf einer willkürlichen Gewichtung der einzelnen Neuronen basiert, mit dem Erwartungswert abgeglichen und die bei dem Vergleich abgeleitete Differenz anschließend als Fehlerwert ausgedrückt. Anhand dieses Fehlerwertes findet im Anschluss eine Fehlerrückführung statt, welche zu einer Anpassung der Gewichte und damit zur Optimierung des KI-Systems führt.

## Barrierefreiheit

Die Barrierefreiheit bezieht sich auf die Anforderung an Systeme künstlicher Intelligenz, unabhängig vom physischen oder mentalen Zustand des jeweiligen Nutzenden bedienbar zu sein. Auf diese Weise soll sichergestellt werden, dass die Gesamtheit der Gesellschaft von den Potentialen solcher Systeme profitieren kann. Voraussetzung dafür ist die Implementierung verschiedener Interaktionsmöglichkeiten. Die Kommunikation kann beispielsweise durch Spracherkennung oder Visualisierung unterstützt werden.

## Big Data

Big Data beschreibt das weltweit exponentiell wachsende Datenaufkommen, dessen Erfassung und Verarbeitung nur noch mit neuartiger und zu diesem Zweck geschaffenen kostenintensiven Technologien realisierbar ist. Für den Umgang mit dem Begriff hat sich eine Charakterisierung entlang der drei V's durchgesetzt:

- ▶ **Volume** – Erfassung und Verarbeitung von enorm großen Datenmengen. Im Verlauf der 2010er Jahre hat sich der

weltweite Datenbestand von 1,2 Zettabyte auf etwa 40 Zettabyte jährlich verdoppelt. Bis zum Jahr 2025 ist mit einem Wachstum auf 175 Zettabyte zu rechnen.

- ▶ **Variety** – Erfassung und Verarbeitung von Daten mit einem unterschiedlichen Grad an Strukturiertheit. Differenziert wird dabei in strukturierte, semi-strukturierte und unstrukturierte Daten.
- ▶ **Velocity** – Erfassung und Verarbeitung von Daten in extrem hoher Geschwindigkeit, teilweise sogar (fast) in Echtzeit.

Im weiteren Zeitverlauf wurden mit den Eigenschaften Veracity und Value zwei weitere V's zur Charakterisierung von Big Data ergänzt, wobei sich Veracity auf die Datenqualität und Value auf die unternehmerische Nutzbarkeit von Big Data bezieht.

## Black Box

Eine Black Box ist ein in sich geschlossenes System, das keine Auskunft darüber gibt, wie es zu einem bestimmten Ergebnis gekommen ist und das selbst für Experten:innen auf Grund der Komplexität nicht mehr interpretierbar ist. Die in diesem Zusammenhang beschriebene Komplexität entsteht durch die vom neuronalen Netz selbstständig vorgenommene und für die Nutzenden nicht mehr nachvollziehbare Gewichtung von einzelnen Merkmalen.

## Bot

Als Bot wird ein Programm bezeichnet, welches wiederkehrende Aufgaben automatisiert bearbeiten kann und damit in verschiedenen Anwendungsszenarien einsetzbar ist. Als bekannteste Programme dieser Art sind Chatbots, Social Bots und Game Bots zu nennen. Mithilfe von Machine-Learning-Methoden

werden die Bots so trainiert, dass sie in der Lage sind, adäquat auf die Anforderungen der zukünftigen Bedienenden einzugehen. Ein Chatbot muss in diesem Zusammenhang beispielsweise auf eine individuelle Kundenanfrage reagieren können.

## Clustering

Clustering bezeichnet ein Verfahren, bei dem die Eingabedaten auf Basis eines Ähnlichkeitsmaßes autonom in Gruppen eingeteilt werden. Auf diesem Weg kann eine Kompression der Informationen erfolgen. Beim Clustering gibt es keine vorher definierten Gruppen. Diese werden vom KI-System erst mithilfe eines Trainingsdatensatzes bestimmt. Grundlage für die Gruppierung stellen durch das KI-System autonom identifizierte Merkmale dar. Entsprechend wird das Verfahren ausschließlich im Rahmen der unüberwachten Entwicklung von Machine-Learning-Modellen eingesetzt.

## Data Bias

Datenbasierte Voreingenommenheit (Data Bias) beschreibt eine Art der Implementierung von Vorurteilen und strukturellen Diskriminierungen aus der realen Welt in KI-basierte Systeme.

Bei dieser Art von Bias werden dem System Vorurteile in der Regel implizit durch ein Modelltraining mit Daten, deren Inhalt strukturelle Benachteiligungen widerspiegelt, angelehrt.

Formen von Diskriminierung, welche dem System in der Regel explizit durch vorurteilsbehaftete Algorithmen angeeignet werden nennt man Algorithmische Voreingenommenheit (Algorithmic Bias).

## Data Mining

Data Mining ist die systematische Analyse von großen Datenmengen (Big Data). Ziel ist es, Muster oder Zusammenhänge in den Daten zu erkennen und diese zu extrahieren.

## Datenschutz

Datenschutz beschreibt den Schutz des Einzelnen vor der missbräuchlichen Verarbeitung seiner personenbezogenen Daten. Auch indirekte personenbezogene Daten fallen unter den Datenschutz. Der rechtliche Umgang mit personenbezogenen Daten ist in der Datenschutzgrundverordnung (DSVGO) und dem Bundesdatenschutzgesetz (BDSG) geregelt.

In Hinblick auf die Thematik künstlicher Intelligenz ergibt sich damit ein Spannungsfeld. So besteht ein Grundsatz der DSGVO in der Minimierung von Datenaufkommen, wohingegen die Qualität von KI-Systemen entscheidend von der Menge an verfügbaren Trainingsdaten abhängig ist. KI-Systeme dürfen personenbezogene Daten nur auf Basis freiwilliger Einwilligung und nur auf einen expliziten Fall bezogen verarbeiten (Art. 9 Abs. 2 DSGVO).

## Deep Learning

Deep Learning, als Teilgebiet des Machine Learnings, umfasst Datenstrukturen zur Verarbeitung und Erkennung von Mustern, deren Aufbau von der neurobiologischen Architektur des menschlichen Gehirns inspiriert ist. Dabei werden Daten in hierarchisch organisierten Ebenen immer weiter abstrahiert, um abschließend relevante Information in den Eingangsdaten angemessen repräsentieren zu können. Die benannten Algorithmik-Strukturen werden entsprechend häufig als künstliche neuronale Netze bezeichnet (KNN).

## Dynamic Pricing

Dynamic Pricing bezeichnet einen Preissetzungsmechanismus, bei dem die Preise für ein Produkt im Tagesverlauf mehrmals angepasst werden. Vor allem im E-Commerce ist diese Art der Preissetzung weit verbreitet (v. a. im Luftverkehr). Bei der Preisgestaltung spielen verschiedene Einflussfaktoren, wie Zeitpunkt, individuelle Merkmale der Kund:innen oder Preise von Wettbewerbern eine Rolle. Dieser Umstand führt dazu, dass zum selben Zeitpunkt zwei Kund:innen unterschiedliche Preise angeboten werden können. Um schnell und effektiv auf die dynamischen Einflussfaktoren reagieren zu können, kommen beim Dynamic Pricing Machine-Learning-Methoden zum Einsatz, um auf diesem Weg die Preisgestaltung zu optimieren.

## Erklärbare KI

Das Forschungsfeld der erklärbaren KI (Englisch: Explainable AI, kurz: XAI) umfasst Prozesse und Methoden, um KI-Systeme so auszustatten, dass den Nutzenden interne Funktionsweisen und Einflussfaktoren einer konkreten Einzelentscheidung aufgezeigt und erklärt werden. Damit die Nutzenden möglichst effizient mit dem KI-System umgehen können, sollten die erklärenden Eigenschaften an das Wissen der Nutzenden angepasst sein.

## Expertensysteme

Expertensysteme, als Teilgebiet der symbolischen KI, setzen sich in Form von explizit hinterlegten Fakten und Regeln mit einem abgegrenzten Fachgebiet auseinander. Probleme aus diesem abgegrenzten Fachgebiet werden in der Folge überwiegend in Form von Wenn-Dann-Funktionen bearbeitet. Die Qualität der Lösungen

ist dabei mit den alternativ von Fachleuten erarbeiteten Lösungsvorschlägen vergleichbar.

## Generative Adversarial Networks (GAN's)

Generative Adversarial Networks (deutsch: Generative gegensätzliche Netzwerke) stellen eine relativ neuartige Methodik des unüberwachten Lernens dar, bei der zwei Netzwerke (Generator und Diskriminator) zueinander in Konkurrenz stehen. Dabei versucht der Generator, basierend auf einem Trainingsdatensatz, weitere möglichst realistische Daten dieser Art zu erzeugen, während der Diskriminator versucht, die synthetisch erzeugten Daten von den realen Werten zu unterscheiden. Mit Hilfe eines Backpropagation-Algorithmus werden die beiden konkurrierenden Netzwerke simultan optimiert.

## Gewichte

Im Zusammenhang mit KI-Systemen bezeichnen Gewichte einen Zahlenwert, der den Verbindungen einzelner Verarbeitungseinheiten (z. B. Neuronen) zugeordnet wird. Die Gewichte unterliegen während der Modellentwicklung einer automatischen und dynamischen Anpassung. Während des Anpassungsprozesses wird das Ziel verfolgt, die entscheidenden Informationen der Eingangssignale mit einer stärkeren Gewichtung in den Entscheidungsprozess einzubeziehen.

## Informationelle Selbstbestimmung

Informationelle Selbstbestimmung bezeichnet das Recht auf Souveränität hinsichtlich der Informationen über die eigene Person. Dabei wird das Ziel verfolgt, den

Einfluss, den Außenstehende auf das eigene Verhalten auswirken können, zu minimieren.

## KI-basierte Geschäftsmodelle

Von einem KI-basierten Geschäftsmodell ist die Rede, wenn in mindestens einem für das Geschäftsmodell relevanten Bereich ein KI-System zum Einsatz kommt. Die Motivation für den Einsatz besteht dabei darin, eine Wettbewerbsposition zu halten oder zu verbessern.

## KI-Ethik

Das Forschungsfeld der KI-Ethik setzt sich mit der Erarbeitung und Identifizierung von (gesellschaftlich anerkannten) Werten, Prinzipien und Techniken als moralische Leitlinie für eine nachhaltige Entwicklung und Nutzung von KI-Systemen auseinander. Teilgebiete der KI-Ethik sind u. a. Maschinenethik, Datenethik sowie das moralische Verhalten von Menschen beim Entwerfen, Programmieren, Verwenden und Behandeln von KI.

## Klassifikation

Klassifikation bezeichnet ein Verfahren, bei dem die Eingabedaten vor dem Lernprozess durch die Entwickler:innen einzelnen Klassen zugeordnet werden und auf diesem Weg vordefinierte Gruppen entstehen. Diese Gruppen (inklusive ihrer definierten Merkmale) dienen anschließend als Trainingsdaten für die Modellentwicklung. Das abgeleitete Modell kann in der Folge für die Bewertung und Einordnung von Daten mit ähnlicher Struktur eingesetzt werden. Entsprechend der im Voraus beschriebenen vordefinierten Gruppen wird das Verfahren im Rahmen der überwachten Entwicklung von Machine-Learning-Modellen eingesetzt.

## Künstliche Intelligenz

Künstliche Intelligenz (KI) soll Aspekte menschlicher Verhaltensweisen in sich nachbilden, um auf diesem Wege „menschlich“ agieren zu können, ohne es zu sein. Hierzu zählen Eigenschaften und Fähigkeiten wie das Lösen von Problemen, Erklären, Lernen, Sprachverstehen und die flexiblen Reaktionen eines Menschen. Dabei wird im Allgemeinen zwischen starker und schwacher KI unterschieden.

## Künstliches Neuronales Netz (KNN)

Künstliche neuronale Netze bestehen entsprechend dem neurobiologischen Vorbild in der Regel aus Eingabe-, Verarbeitungs- und Ausgabeschichten und stellen eine etablierte Struktur für die Entwicklung von KI-Systemen dar. Die beschriebenen Schichten setzen sich dabei aus jeweils einzelnen Neuronen zusammen. Mit der Ergänzung von vielzähligen Verarbeitungsschichten (sog. „Hidden Layers“), welche der Abstraktion von Eingabedaten dienen, entwickelten sich künstliche neuronale Netze zur Grundlage für Deep-Learning-Verfahren.

## Lernen

Lernen beschreibt den Prozess, wie ein Vorgang oder eine Aufgabe bei Wiederholung ausgeführt oder angepasst wird. Die Veränderung basiert auf Erfahrungen und Erlebnissen, die Lernende in der Vergangenheit erlebt haben sowie deren Projektion auf Gegenwart und Zukunft. Dadurch können Lernende auf neue Begebenheiten und Umweltbedingungen reagieren und sich daran anpassen. Auch Algorithmen sind in der Lage, zu lernen. Mithilfe von Daten aus der Vergangenheit können diese Rückschlüsse auf

die Gegenwart bzw. Zukunft ziehen und Aussagen über Sachverhalte und Entscheidungen treffen.

## Machine Learning

Machine Learning bezeichnet die Entwicklung von Algorithmen, die es Computerprogrammen ermöglichen, sich automatisch auf Basis von Erfahrungswerten zu optimieren. Durch diesen Bottom-up-Ansatz, bei dem Wissen auf Basis von Erfahrung entsteht, stellt Machine Learning ein Teil der Subsymbolischen KI dar. Das Wissen wird dabei in der Regel durch Methoden des überwachten, unüberwachten oder verstärkenden Lernens generiert. Das daraus resultierende Modell kann anschließend auf vorher unbekannte Daten angewandt und zu deren Interpretation genutzt werden.

## Natural Language Processing

Natural Language Processing (NLP) beschreibt computergestützte Techniken zur maschinellen Erkennung und Verarbeitung von natürlicher Sprache. Das Ziel ist dabei, die direkte Kommunikation zwischen Mensch und Computer auf Basis natürlicher Sprache zu ermöglichen, die zwischenmenschliche Kommunikation durch maschinelle Übersetzung zu erleichtern und Texte in natürlicher Sprache maschinell analysieren zu können. NLP verwendet hierfür Erkenntnisse der Sprachwissenschaften sowie Methoden und Techniken der künstlichen Intelligenz, speziell aus dem Bereich Machine Learning und der Subkategorie Deep Learning.

## Neuron

Das Neuron stellt die elementare Verbindungseinheit eines künstlichen neuronalen Netzes dar.

Die Summe aller Neuronen ist im Stande, Eingangssignale parallel zu verarbeiten und dient damit der effizienten Generierung von Gewichten innerhalb des künstlichen neuronalen Netzes.

## Resilienz

Resilienz beschreibt die Eigenschaft informationstechnischer Systeme, eine Funktionsstörung eigenständig und ohne schwerwiegende Folgen zu bewältigen.

## Robotik

Die Robotik beschäftigt sich mit der Entwicklung und Herstellung von Robotern. Ziel ist es, dass Roboter dem Menschen Teile seiner Arbeit abnehmen. Neben körperlichen Tätigkeiten (Industrieroboter) werden auch immer mehr kognitive Tätigkeiten (Serviceroboter) übernommen. Die Robotik ist ein interdisziplinäres Feld aus Elektrotechnik, Mechanik und künstlicher Intelligenz.

## Robustheit

Robustheit beschreibt die Eigenschaft informationstechnischer Systeme, bei böartigen Angriffen oder zufällig herbeigeführten Fehlern trotzdem eine Funktionsfähigkeit zu gewährleisten, die dem Normalbetrieb gleicht oder diesem nur minimal vermindert gegenübersteht.

## Schwache KI

Schwache KI bezeichnet intelligente Systeme, die trainiert und angelernt werden, um bestimmte Aufgaben zu erledigen. Dabei werden sie für die Ausführung in einem definierten Aufgabenbereich nicht explizit programmiert. Zum aktuellen Zeitpunkt sind alle jemals entwickelten KI-Systeme dieser Kategorie zuzuordnen.

## Smart Data

Smart Data sind Datensätze, die aus Big Data gewonnen werden. Diese Daten sind strukturiert und qualitativ hochwertig, sodass diese für den weiteren Erkenntnisgewinn genutzt werden können.

## Starke KI

Starke KI kann von der Leistungsfähigkeit her mit menschlichen Fähigkeiten verglichen werden und wird nicht auf ein definiertes Aufgabengebiet limitiert. Problemlösung, abstraktes Denken oder Ideenfindung wären hiermit möglich.

## Subsymbolische KI

Die Subsymbolische KI umfasst alle KI-Systeme, deren Funktionsfähigkeit auf einem Bottom-up-Ansatz beruhen. Entsprechend basieren die von KI-Systemen dieser Art ausgegebenen Erkenntnisse auf selbstständig geschlussfolgerten Mustern und Wissen aus Beispieldaten. Ein Vorteil dieses Ansatzes besteht in der meist hohen Anpassungsfähigkeit aufgrund der flexiblen Struktur zur Wissenserfassung und Ergebniserzeugung. Der daraus ebenfalls resultierende geminderte Grad an realisierbarer Transparenz stellt dagegen einen nicht zu vernachlässigenden Nachteil solcher Systeme dar.

## Symbolische KI

Die symbolische KI umfasst alle KI-Systeme, deren Funktionsfähigkeit darauf basiert, dass die von ihnen ausgegebenen Erkenntnisse auf explizit vorgegebenen Wissensrepräsentationen beruhen. Aufgrund der klar vorgegebenen Strukturen besteht ein Vorteil dieses Ansatzes darin, einen hohen Grad an Transparenz zu erreichen. Das daraus ebenfalls resultierende geminderte Anpassungsvermögen ist dagegen nachteilig.

## Testdatensatz

Ein Testdatensatz ist ein Datensatz, welcher für die Modellbewertung von KI-Systemen, insbesondere im Bereich der überwachten Lernverfahren, dient. Der Testdatensatz ist vor allem dafür geeignet, das auf Basis der Trainingsdaten entwickelte Modell auf Über- oder Unteranpassung zu überprüfen.

## Trainingsdatensatz

Ein Trainingsdatensatz ist ein Datensatz, welcher für die Modellentwicklung von KI-Systemen, insbesondere im Bereich der überwachten Lernverfahren, dient. Die Qualität der Trainingsdaten und der darin enthaltenen Muster ist dabei von entscheidender Bedeutung für die spätere Performance des jeweiligen KI-Systems.

## Transparenz

Ein KI-System kann als transparent bezeichnet werden, wenn ein außenstehender Dritter für die internen Vorgänge aufgrund von begrenzter Komplexität ein Verständnis entwickeln kann und das Verhalten der Anwendung damit vollständig nachvollziehbar ist. Auch die Systemumgebung sollte transparent gestaltet sein, indem Nutzer:innen u. a. kenntlich gemacht wird, dass es sich um eine Interaktion mit einem KI-System handelt. Darüber hinaus sollte deutlich gemacht werden, zu welchem Zweck und in welchem Umfang für die beschriebene Interaktion Nutzerdaten erfasst und verarbeitet werden.

## Turing-Test

Der Turing-Test ist benannt nach dem englischen Informatiker und Kryptoanalytiker Alan Turing. Er ist der weltweit bekannteste Test zur Evaluation von KI-Systemen

und die Basis für Diskussionen rund um das Thema der künstlichen Intelligenz. Bei diesem Test kommunizieren die Nutzenden per Maus und Tastatur mit zwei Systemen und versuchen herauszufinden, in welchem Fall mit einem Computer und in welchem Fall mit einem Menschen kommuniziert wurde.

## Überanpassung

Überanpassung (Englisch: Overfit) beschreibt die Eigenschaft eines KI-Systems, welches im Rahmen der Lernphase zu sehr auf die Inhalte der Trainingsdaten abgestimmt wurde, indem es ab einem bestimmten Zeitpunkt für eine Generalisierung irrelevante Merkmale in den Entscheidungsprozess einbezieht. Entsprechend unzufriedenstellend fällt die Leistung aus, wenn das System vermeintlich ähnliche Testdaten verarbeitet, welche diese irrelevanten Merkmale nicht enthalten. Man spricht davon, dass KI-Systeme mit dieser Eigenschaft ein geschwächtes Generalisierungsvermögen besitzen.

## Überwachtes Lernen

Das überwachte Lernen stellt ein Lernverfahren dar, welches für die Entwicklung von Modellen, sowohl aus dem Bereich der symbolischen als auch der subsymbolischen KI eingesetzt wird. Voraussetzung dafür ist, dass den Trainingsdaten erwartbare Ergebnisse anhand von Labels oder ähnlichen Kategorisierungsmöglichkeiten zugeordnet werden. Auf diesem Weg können vom KI-System vorgenommene Zuordnungen autonom überprüft werden und das Modell anschließend durch passende Algorithmen (u. a. Klassifikation) angepasst werden. Dabei wird das Ziel verfolgt, relevante Muster zu identifizieren und die Erkenntnisse anschließend auf unbekannte Daten anzuwenden.

## Unteranpassung

Unteranpassung (Englisch: Underfit) beschreibt die Eigenschaft eines KI-Systems, welches bereits bei der Anwendung der Trainingsdaten unzufriedenstellende Ergebnisse liefert, da aufgrund mangelnder Spezialisierung des Modells noch keine relevanten Merkmale identifiziert werden können.

## Unüberwachtes Lernen

Das unüberwachte Lernen stellt ein Lernverfahren dar, welches für die Entwicklung von Modellen, sowohl aus dem Bereich der symbolischen als auch der subsymbolischen KI eingesetzt wird. Für die Modellentwicklung stehen entgegengesetzt zum überwachten Lernen allerdings nur nicht-kategorisierte Trainingsdaten zur Verfügung. Häufig wird das unüberwachte Lernen eingesetzt, um durch geeignete Verfahren (z. B. Clustering) und ohne Feedback von außen, Muster und Merkmale in den Daten zu identifizieren, die im Rahmen einer konventionellen Analyse eventuell unentdeckt geblieben wären.

## Validierungsdatensatz

Ein Validierungsdatensatz ist ein Datensatz, welcher insbesondere im Bereich der überwachten Lernverfahren zur Optimierung der Hyperparameter (Parameter, der vor Beginn des Lernprozesses festgelegt wird) dient. Durch Optimierung der Hyperparameter kann dem Zustand einer Überanpassung vorgebeugt werden.

## Vertrauensvolle und verantwortungsvolle KI

Vertrauensvolle und verantwortungsvolle KI (Englisch: responsible and trustworthy AI) stellen aktuell relevante Konzepte für

eine nachhaltige Entwicklung von künstlicher Intelligenz dar. Inhaltlich umfassen beide Konzepte weitgehend identische Charakteristiken, welche sicherstellen sollen, dass das Feld der künstlichen Intelligenz in Zukunft stets den fundamentalen Menschenrechten und Werten gerecht wird. Die angesprochene Verantwortung soll und kann dabei nicht mit informationstechnischen Mitteln an das System übertragen werden, sondern liegt in den Händen aller am Prozess der Entwicklung und Anwendung des Systems beteiligten Menschen.

Folgende Charakteristiken sind dabei unter anderem von Relevanz: Barrierefreiheit, Informationelle Selbstbestimmung, Robustheit, Resilienz, Transparenz, Erklärbarkeit (Erklärbare KI).

## Verstärkendes Lernen

Das verstärkende Lernen stellt ein Lernverfahren dar, welches für die Entwicklung von Machine-Learning-Modellen eingesetzt wird. Mittels einer zu definierenden Belohnungsfunktion werden dem System auf anfangs noch zufällige Aktionen je nach Erfolg positives oder negatives Feedback gespiegelt. Durch das Streben des Systems nach der größtmöglichen Belohnung nähert sich der Algorithmus allmählich der bestmöglichen Lösung für das gegebene Problem. Für den gesamten Prozess wird kein vorheriges Datenmaterial benötigt, wodurch sich dieses Lernverfahren vom überwachten und unüberwachten Lernen abgrenzt.

## Literaturverzeichnis

Abdelkafi, D. h., Döbel, I., Drzewiecki, J. D., Meironke, A., Niekler, A., & Ries, S. (2019). Künstliche Intelligenz (KI) im Unternehmenskontext. Leipzig.

Arrieta, A. B., Díaz-Rodríguez, N., Ser, J. D., & Bennetot, A. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, 58, 82-115.

Bissantz, N., & Hagedorn, J. (2009). Data Mining (Datenmustererkennung). *Wirtschaftsinformatik*, 51(1), 139-144.

Buxmann, P., & Schmidt, P. (2019). *Künstliche Intelligenz*. Springer Berlin Heidelberg.

Dignum, V. (2019). *Responsible artificial intelligence: how to develop and use AI in a responsible way*. Springer Nature.

Döbel, I., Leis, D. M., Vogelsang, M. M., & Neustroev, D. (2018). Maschinelles Lernen - Kompetenzen, Anwendungen und Forschungsbedarf. *Sankt Augustin: Fraunhofer-Gesellschaft (IAS, IMW, Zentrale)*.

Gentsch, P. (2017). *Künstliche Intelligenz für Sales, Marketing und Service: Mit AI und Bots zu einem Algorithmic Business – Konzepte, Technologien und Best Practices*. Springer-Verlag.

Grigoleit, D. S. (2019). Natural Language Processing. *Europäische Sicherheit & Technik*, 69.

Grum, Sultanow, Friedmann, Ullrich, & Gronau. (2021). Symbolische KI. In *Tools des Maschinellen Lernens*. GITO.

Ishwarappa, & Anuradha, J. (2015). A Brief Introduction on Big Data 5Vs Characteristics and Hadoop Technology. *Procedia Computer Science*, 48, 319-324.

Jaume-Palásí, L., & Spielkamp, M. (2017). Ethik und algorithmische Prozesse zur Entscheidungsfindung oder -vorbereitung. *AlgorithmWatch Arbeitspapier*, 4.

Knoll, A., Burgard, W., & Christaller, T. (2010). Robotik. In G. Görz, & J. Schneeberger (Hrsg.), *Handbuch der künstlichen Intelligenz* (S. 871-930). Walter de Gruyter.

Leslie, D. D. (2019). Understanding artificial intelligence ethics and safety: A guide for the responsible design and implementation of AI.

Meier, A. (2018). Was heißt Big Data? In *Werkzeuge der digitalen Wirtschaft: Big Data, NoSQL & Co.* (S. 5-7). Wiesbaden: Springer Vieweg.

Narahari, Y., Raju, C. V., Ravikumar, K., & Shah, S. (2005). Dynamic pricing models for electronic business. *sadhana*, 30(2), 231-256.

Wahlster, W. (2017). Künstliche Intelligenz als Grundlage autonomer Systeme. *Informatik-Spektrum*, 40(5), 409-418.

Wennker, P. (2020). *Künstliche Intelligenz in der Praxis*. Springer Books.

Witt, B. (2010). *Datenschutz kompakt und verständlich*. Wiesbaden: Vieweg+ Teubner.

Wittpahl, V. (2019). *Künstliche Intelligenz - Technologie | Anwendung | Gesellschaft*. Springer Vieweg.

### Impressum:

Autor: Nick Harnau, Lorenz Laderick  
Redaktion: Sarah Kilz, Matthias Welker  
Grafik: Nick Harnau

Mittelstand 4.0-Kompetenzzentrum eStandards  
Offene Werkstatt Leipzig  
c/o Fraunhofer-Zentrum für Internationales Management und  
Wissensökonomie IMW

### Kontakt:

Tel: +49 341 231039 122  
leipzig@kompetenzzentrum-estandards.digital  
[www.kompetenzzentrum-estandards.digital](http://www.kompetenzzentrum-estandards.digital)

Das Mittelstand 4.0-Kompetenzzentrum eStandards gehört zu Mittelstand-Digital. Mit Mittelstand-Digital unterstützt das Bundesministerium für Wirtschaft und Klimaschutz die Digitalisierung in kleinen und mittleren Unternehmen und dem Handwerk.

### 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 Klimaschutz ermöglicht die kostenfreie Nutzung aller Angebote von Mittelstand-Digital.

Weitere Informationen finden Sie unter  
[www.mittelstand-digital.de](http://www.mittelstand-digital.de)