

LEHRERHANDBUCH

Gesellschaftliche und ethische Aspekte von KI

Thema	Klärung des Begriffs „KI“
Zielgruppe	ab Klasse 10, besonders geeignet für die Oberstufe
Fachbezüge	Gemeinschaftskunde, Ethik, Medienbildung, Informatik
Zeitrahmen	45–90 Minuten
Material	Arbeitsblatt 1 Webcode: https://box4bot.de/webcode/buch_kig.html 1-K4MN3P (PDF) 1-K4MN3D (DOCX)

Didaktische Funktion

Dieses erste Kapitel soll den Begriff KI öffnen, ohne bereits zu stark in technische Details einzusteigen. Entscheidend ist nicht, eine perfekte Definition zu vermitteln. Ziel ist vielmehr, vorschnelle Vorstellungen zu korrigieren: KI ist weder Magie noch ein denkendes Wesen, aber auch nicht einfach „normale Software“. Zugleich sollen die Lernenden erste Kriterien für eine kritische Bewertung von KI-Systemen entwickeln: Welche Daten werden genutzt, welches Ziel wird verfolgt, welche Fehler können entstehen und wer trägt Verantwortung?

Die Schülerinnen und Schüler sollen erkennen, dass KI ein Sammelbegriff ist. Systeme können nach festen Regeln arbeiten, Suchräume durchprobieren, Muster aus Beispielen lernen oder neue Inhalte erzeugen. In realen Anwendungen überschneiden sich diese Formen häufig.

Der Baustein bereitet die späteren Kapitel vor. Besonders wichtig sind drei Einsichten: Erstens ist nicht jedes digitale System KI. Zweitens hängen viele moderne KI-Systeme stark von Daten und Beispielen ab. Drittens lautet die entscheidende Frage nicht nur „Ist das eine KI?“, sondern „Was tut das System, mit welchen Daten, welchem Ziel und welchen möglichen Fehlern?“

Warum dieses Kapitel für das Verständnis wichtig ist

Dieses Kapitel ist für das Funktionsverständnis zentral, weil es die wichtigste begriffliche Fehlspur früh korrigiert: KI wird im Alltag oft so beschrieben, als handle es sich um eine Art künstliches Subjekt. Schülerinnen und Schüler begegnen Chatbots, Empfehlungssystemen oder Bildgeneratoren jedoch meist über ihre Oberfläche. Diese Oberfläche wirkt dialogisch, kreativ oder sogar verständnisvoll. Ohne begriffliche Klärung entsteht leicht der Eindruck, dass hinter der Antwort ein denkendes Gegenüber steht. Gerade deshalb muss der Unterricht früh zwischen Oberfläche und Funktionsweise unterscheiden. Die Oberfläche zeigt, was ein System ausgibt: eine Antwort, eine Empfehlung, eine Bewertung, ein Bild. Die Funktionsweise fragt dagegen, wie diese Ausgabe entsteht: durch Daten, Modellstruktur, Zielvorgaben, Wahrscheinlichkeiten, Klassifikation oder Generierung. Diese Unterscheidung ist die Grundlage für kritische Perspektiven. Wer nur die Oberfläche beurteilt, überschätzt leicht Verstehen, Objektivität oder Zuverlässigkeit. Wer die Funktionsweise mitdenkt, kann gezielter fragen, wo Grenzen, Bias, Fehlannahmen und Verantwortungsprobleme entstehen.

Das Kapitel führt außerdem eine zweite Grundunterscheidung ein: Nicht jede Automatisierung ist KI. Ein Taschenrechner rechnet schneller als ein Mensch, aber er lernt nicht aus Beispielen. Eine Waschmaschine mit Programmwahl kann automatisch arbeiten, ohne Muster aus Daten zu erkennen. Ein Spamfilter oder eine Bilderkennung dagegen arbeitet typischerweise mit gelernten Mustern. Diese Unterscheidung ist wichtig, weil spätere Kapitel auf ihr aufbauen. Datenqualität, Bias, Halluzinationen und Fehler bei Klassifikationen lassen sich nur verstehen, wenn klar ist, dass viele KI-Systeme ihr Verhalten aus Beispieldaten ableiten.

Die vier Grundideen – Regeln ausführen, suchen und planen, Muster erkennen, Inhalte erzeugen – bilden ein erstes mentales Ordnungssystem. Es ist bewusst vereinfacht, aber didaktisch tragfähig. Schülerinnen

und Schüler brauchen erst einmal keine Detailkenntnisse über Transformer, Backpropagation oder Embeddings (Wenn Sie diese Begriffe neugierig machen, fragen Sie die KI Ihres Vertrauens). Sie brauchen ein Raster, mit dem sie verschiedene Systeme vergleichen können: Arbeitet das System nach festen Regeln? Prüft es viele Möglichkeiten? Ordnet es Eingaben Klassen zu? Erzeugt es neue Ausgaben? Diese Fragen helfen, technische und gesellschaftliche Debatten zu entmystifizieren.

Besonders wichtig ist die letzte Leitfrage des Arbeitsblatts: „Was tut das System genau?“ Sie verschiebt den Blick weg vom Etikett „KI“ hin zur Funktionsweise. Dadurch wird später verständlich, warum ein Empfehlungssystem andere Probleme erzeugt als ein Bildklassifikator, warum ein Sprachmodell anders geprüft werden muss als eine Suchmaschine und warum ein Scoring-System andere Verantwortungsfragen aufwirft.

→ **Grundprinzipien des Bandes (Bandeinleitung)**: Dieses Kapitel knüpft besonders an **Grundprinzip 1** an („KI-Systeme verstehen ihren Output nicht“). Alle fünf Grundprinzipien sind in der *Bandeinleitung für Lehrkräfte* ausführlich erläutert.

Lernziele

Die Schülerinnen und Schüler

- unterscheiden einfache regelbasierte Systeme von datengetriebenen KI-Systemen,
- beschreiben KI als Sammelbegriff für verschiedene technische Verfahren,
- nennen zentrale Grundideen: Regeln, Suche, Mustererkennung, Generierung,
- erklären, warum KI-Systeme nicht wie Menschen verstehen oder verantwortlich handeln,
- benennen typische Grenzen und Fehlerquellen von KI-Systemen,
- beurteilen KI-Anwendungen im Hinblick auf Nutzen, Risiken und Verantwortung,
- wenden die Begriffe auf alltagsnahe Beispiele an,
- formulieren eine erste eigene Arbeitsdefinition von KI.

Hinweise zur Durchführung

Der Warm-up sollte bewusst ohne vorherige Definition beginnen. Es ist erwünscht, dass Schülerinnen und Schüler uneinheitlich entscheiden. Gerade Grenzfälle wie Navigations-App, Schachprogramm oder Waschmaschine zeigen, dass der KI-Begriff im Alltag nicht trennscharf verwendet wird.

Bei der Besprechung sollte nicht vorschnell korrigiert werden. Die Einordnung von Grenzfällen ist nicht vor allem eine Frage von „richtig“ oder „falsch“, sondern eine Frage der Begründung. Gerade an den Begründungen wird KI-Literacy sichtbar: Lernende zeigen, ob sie nur ein Etikett vergeben oder ob sie Daten, Verfahren, Ziel und mögliche Folgen mitdenken. Besser ist deshalb eine Sortierung nach Begründungen:

- „Das System folgt nur festen Regeln.“
- „Das System passt sich an Daten an.“
- „Das System erkennt Muster.“
- „Das System erzeugt neue Inhalte.“
- „Das System wirkt intelligent, versteht aber nicht.“

Die Lehrkraft sollte Begriffe wie Daten, Training, Prognose, Bias und Verantwortlichkeit an geeigneten Stellen behutsam einführen. Es geht nicht darum, alle Begriffe sofort vollständig zu erklären. Sie sollen als Denksuren sichtbar werden: KI-Systeme nutzen Daten, werden häufig trainiert, geben Prognosen oder Zuordnungen aus, können verzerrt sein und müssen verantwortet eingesetzt werden.

Die kurze Entwicklungslinie ist kein Geschichtskapitel. Sie soll zeigen, dass KI mehrere Entwicklungsstufen durchlaufen hat: frühe Dialogsimulation, spezialisierte Spielsysteme, datengetriebene Bilderkennung, generative Sprachmodelle. Wichtig ist dabei die Abgrenzung zwischen Wirkung und Funktionsweise. ELIZA wirkte gesprächsfähig, verstand aber nicht. Deep Blue war extrem stark im Schach, aber kein allgemeines Denkmodell. Moderne Sprachmodelle erzeugen eindrucksvolle Texte, bleiben aber daten- und modellgebundene Systeme (und können nicht gut Schach spielen).

Lösungshinweise zum Warm-up

Die Tabelle ist bewusst offen. Einige sinnvolle Einordnungen:

System	Mögliche Einordnung	Begründung
Taschenrechner	eher keine KI	führt feste Rechenregeln aus
Navigations-App	Grenzfall / teilweise KI	Routenplanung regel- und suchbasiert; moderne Systeme nutzen Verkehrsdaten und Prognosen
Spamfilter	KI	erkennt Muster in Nachrichten und lernt aus Beispielen
Schachprogramm	Grenzfall bis KI	klassische Systeme: Suche und Bewertung; moderne Systeme oft auch datengetrieben
Gesichtserkennung	KI	erkennt Muster in Bilddaten
Chatbot	KI	erzeugt Sprache nach gelernten Mustern
Online-Shop-Empfehlung	KI	nutzt Nutzungsdaten und Muster, um Produkte zu ranken
Waschmaschine mit Automatikprogramm	eher keine KI / Grenzfall	feste Programme; bei Sensorik und Anpassung Grenzfall, aber meist kein Lernen aus großen Beispieldaten

Wichtig ist, dass Schülerinnen und Schüler die Einordnung begründen. Die Begründung ist fachlich wertvoller als das Etikett.

Lösungshinweise zu Teil 3

Ein Schachprogramm bewertet mögliche Züge. Vor allem: Suchen und planen. Klassische Schachprogramme durchsuchen viele mögliche Zugfolgen und bewerten Stellungen nach festgelegten Kriterien. Je nach System können zusätzlich Mustererkennung oder gelernte Bewertungsfunktionen beteiligt sein; deshalb ist „Schachprogramm“ keine einheitliche technische Kategorie.

Ein Spamfilter sortiert E-Mails aus. Vor allem: Muster erkennen. Er nutzt Merkmale von Nachrichten und entscheidet, ob sie wahrscheinlich unerwünscht sind. Die Grenze liegt darin, dass er nicht „weiß“, ob eine Nachricht wirklich unerwünscht ist; er trifft eine Zuordnung auf Basis gelernter Hinweise.

Ein Bildgenerator erstellt ein Bild zu einer Beschreibung. Vor allem: Inhalte erzeugen. Zugleich nutzt er gelernte Muster aus Trainingsdaten. Die Grenze liegt darin, dass das erzeugte Bild plausibel zur Beschreibung passen kann, ohne eine reale Szene oder eine überprüfte Aussage abzubilden.

Eine Navigations-App berechnet eine Route. Vor allem: Suchen und planen. Moderne Systeme können zusätzlich Verkehrsdaten und Muster aus früheren Situationen nutzen. Die Grenze liegt darin, dass die App eine optimierte Route vorschlägt, aber nicht alle menschlichen Gründe kennt, etwa Unsicherheit, Barrierefreiheit oder persönliche Vorlieben.

Ein Sprachmodell schreibt eine Erklärung. Vor allem: Inhalte erzeugen. Grundlage sind gelernte Muster aus Textdaten. Die Grenze liegt darin, dass eine Erklärung sprachlich überzeugend wirken kann, ohne dass jede Aussage richtig, vollständig oder belegt ist.

Eine App erkennt Gesichter auf Fotos. Vor allem: Muster erkennen. Sie ordnet Bildmerkmale einer Kategorie oder Person zu. Die Grenze liegt darin, dass Erkennungsleistung von Trainingsdaten, Bildqualität, Licht, Alter, Hauttönen und Einsatzkontext abhängen kann.

Eine Lernplattform empfiehlt Aufgaben. Vor allem: Muster erkennen und sortieren. Sie nutzt Daten über Bearbeitung, Fehler, Tempo oder bisherige Leistungen, um passende Aufgaben vorzuschlagen. Die Grenze liegt darin, dass solche Daten Lernverhalten nur ausschnittsweise zeigen und leicht mit Motivation, Sprache, Technikzugang oder Arbeitsbedingungen verwechselt werden können.

Hinweis zu Bias

Bei einigen Beispielen kann bereits der Begriff Bias eingeführt werden. Bias meint hier eine systematische Verzerrung, nicht einfach einen einzelnen Fehler. Ein Spamfilter kann bestimmte Dialekte, Schreibweisen oder Sprachen häufiger falsch einstufen, wenn sie in den Trainingsdaten unzureichend vertreten waren. Gesichtserkennung kann bei bestimmten Hauttönen, Altersgruppen oder Aufnahmebedingungen schlechter funktionieren, wenn die Trainings- und Testdaten nicht vielfältig genug waren. Eine Lernplattform kann Lernende falsch einschätzen, wenn sie nur Bearbeitungszeit oder Klickverhalten auswertet und nicht versteht, warum jemand langsam arbeitet.

Für dieses erste Kapitel genügt eine knappe Klärung: Bias entsteht nicht, weil die KI „Vorurteile hat“ wie ein Mensch, sondern weil Daten, Zielgrößen oder Einsatzbedingungen systematisch verzerrt sein können. Die genaue Analyse folgt in einem späteren Kapitel.

Hinweise zur Zusatzaufgabe

Die Zusatzaufgabe verbindet Begriffsarbeit mit kritischer Bewertung. Sie ist besonders geeignet, um sichtbar zu machen, ob Lernende KI-Systeme nicht nur benennen, sondern funktional untersuchen können. Gute Antworten enthalten keine perfekte technische Analyse, aber sie unterscheiden mindestens vier Ebenen:

- Daten: Welche Eingaben, Nutzungsdaten, Bilder, Texte oder Verhaltensspuren könnten verwendet werden?
- Nutzen: Welches Problem soll das System lösen oder welche Aufgabe erleichtert es?
- Fehler oder Risiken: Welche Fehlannahmen, falschen Zuordnungen, Verzerrungen oder Abhängigkeiten können entstehen?
- Verantwortung: Wer entscheidet über Einsatz, Kontrolle, Korrektur und Folgen?

Bei der Besprechung sollte die Lehrkraft auf unklare Formulierungen achten. „Die KI entscheidet“ sollte präzisiert werden zu: „Das System gibt eine Empfehlung, Sortierung, Klassifikation oder Prognose aus; Menschen oder Institutionen entscheiden, wie sie damit umgehen.“

Hinweise zu typischen Missverständnissen

„KI entscheidet selbst.“ Besser: KI-Systeme erzeugen Ausgaben nach Daten, Modellstruktur und Zielvorgaben. Die Verantwortung für Einsatz und Folgen liegt bei Menschen und Institutionen.

„KI versteht Sprache wie ein Mensch.“ Besser: Sprachmodelle verarbeiten und erzeugen Text anhand gelernter Muster. Sie können sinnvolle Antworten formulieren, ohne menschliches Verständnis zu besitzen.

„Alles Digitale ist KI.“ Besser: Viele digitale Systeme arbeiten ohne Lernen aus Daten. Ein Taschenrechner ist nicht KI, nur weil er schneller rechnet als ein Mensch.

„KI ist immer lernfähig.“ Besser: Viele KI-Systeme wurden in einer Trainingsphase angepasst, lernen aber während der normalen Nutzung nicht automatisch weiter. Diese Unterscheidung wird im Kapitel „Training und Inferenz“ vertieft.

Mögliche 30-Minuten-Fassung

Phase	Min.	Inhalt	Hinweise / Material
Einstieg	5 Min.	Warm-up-Tabelle ausfüllen	Einzelarbeit oder Partnerarbeit
Auswertung	8 Min.	zwei Grenzfälle diskutieren	Navigations-App und Waschmaschine eignen sich besonders
Erarbeitung	8 Min.	Teil 1 und Entwicklungslinie lesen	Fokus auf vier Grundideen

Anwendung	7 Min.	Teil 3 an zwei Beispielen bearbeiten	keine vollständige Tabelle nötig; Zusatzaufgabe nur mündlich anreißen
Abschluss	2 Min.	Mini-Fazit und Rückblick	ein Satz im Heft genügt

Mögliche 45-Minuten-Fassung

Phase	Min.	Inhalt	Hinweise / Material
Einstieg	5 Min.	Warm-up	ohne Definition beginnen
Diskussion	10 Min.	Grenzfälle sortieren	Begründungen sammeln
Erarbeitung	10 Min.	Teil 1 und Teil 2	kurze Entwicklungslinie nicht ausweiten
Anwendung	15 Min.	Teil 3 und Zusatzaufgabe in Partnerarbeit	zwei Beispiele ausführlich erklären lassen; Alltagssystem kurz prüfen
Sicherung	5 Min.	eigene Arbeitsdefinition und Rückblick formulieren	mehrere Definitionen vergleichen

Übergang zu Kapitel 2

Am Ende sollte auf die nächste Leitfrage übergeleitet werden:

Wenn viele KI-Systeme aus Beispielen lernen, dann wird entscheidend, welche Beispiele sie bekommen. Genau deshalb beschäftigt sich das nächste Kapitel mit Daten: Sie sind die Grundlage vieler KI-Systeme, aber sie sind nie einfach „die Wirklichkeit“.