

# AUTONOMIE

*Einschränkung oder Verlust  
der Entscheidungsfreiheit  
oder Kontrolle einer Person,  
Gruppe oder Organisation.*



# KÖRPER & MATERIE

*Physische Verletzung einer  
Person oder Gruppe oder  
Schäden an materiellen  
Gütern.*





# PSYCHE

*Beeinträchtigung der  
emotionalen oder  
psychischen Gesundheit von  
Einzelpersonen, Gruppen  
oder der Gesellschaft.*

# REPUTATION

*Schädigung des Rufs einer  
Person, Gruppe oder  
Organisation.*



# FINANZEN

*Finanzielle Schäden für  
Einzelpersonen,  
Unternehmen oder  
Organisationen.*



# RECHTE & FREIHEITEN

*Gefährdung oder  
Einschränkung  
grundlegender Rechte und  
Freiheiten.*



# GESELLSCHAFT & WIRTSCHAFT

*Beeinträchtigung  
gesellschaftlicher  
Strukturen, Gemeinschaften  
oder Volkswirtschaften.*



# POLITIK

*Manipulation politischer Meinungen, Schädigung politischer Institutionen oder Beeinträchtigung staatlicher Dienstleistungen.*



# UMWELT

*Schädigung oder Zerstörung  
der Umwelt.*



# VERZERRUNG

*Daten enthalten Verzerrungen, die beispielsweise durch gesellschaftliche, historische oder methodische Faktoren entstanden sind.*

*Ein Buchungsdatensatz beinhaltet vorwiegend Daten über Geschäftsreisende. Die Nachfrage nach Reisen oder Nächtigungen am Wochenende ist unterrepräsentiert.*



# UNGLEICHGEWICHT

*Bestimmte Gruppen sind in Datensätzen unterrepräsentiert, andere überrepräsentiert.*

*Ein Datensatz zur Analyse von Kaufverhalten enthält überwiegend Transaktionen von Großkund\*innen, während das Einkaufsverhalten von Privatpersonen kaum repräsentiert ist.*



# HOMOGENITÄT

*Daten decken nicht die gesamte Vielfalt  
der realen Welt ab.*

*Ein Portrait-Datensatz enthält nur Bilder  
von jungen, hellhäutigen Männern.*



# LATENTE MUSTER

*Daten enthalten Merkmale, die mit sensiblen Faktoren korrelieren.*

*Ein Datensatz über Kunden enthält Informationen zu Wohnorten und Einkaufsverhalten, wodurch auch Rückschlüsse auf das Bildungsniveau ermöglicht werden.*



# FREMDZUGRIFF

*Daten sind nicht ausreichend vor unbefugtem Zugriff oder Manipulation geschützt.*

*Ein Datensatz mit Bauplänen wird ungesichert in einer Cloud abgelegt, sodass Unbefugte sensible Informationen einsehen oder verändern können.*



# CONCEPT DRIFT

*Die Eigenschaften von Daten verändern sich über die Zeit, sodass Modelle ohne Update nicht mehr zuverlässig funktionieren.*

*Eine KI zur Erkennung von Spam-Mails wurde mit Altdaten trainiert und erkennt deshalb moderne Betrugsmails nicht mehr.*



# URSPRUNGSRÄTSEL

*Die Herkunft und Erhebungsmethoden von Daten sind unbekannt.*

*Ein Datensatz mit medizinischen Diagnosen stammt aus verschiedensten Kliniken mit unterschiedlichen Erfassungsmethoden, ohne klare Dokumentation der Qualitätsstandards.*



# BIASBLINDHEIT

*Beim Sammeln, Verarbeiten oder Bereinigen der Daten wird Fairness nicht aktiv berücksichtigt.*

*Ein KI-System zur Bewertung von Schüleraufsätzen wurde nur mit Standardsprachgebrauch trainiert und benachteiligt Schüler\*innen mit Migrationshintergrund oder Dialekteinflüssen.*



# LABELCHAOS

*Labels oder Kategorien in den Daten  
enthalten Fehler oder falsche  
Bewertungen.*

*Ein Produktdatensatz mit  
Weihnachtsrabatten enthält fehlerhafte  
Zuordnungen, sodass Artikelrabatte an  
der Kassa oder im Onlineshop nicht  
korrekt berechnet werden.*



# GESETZESVORGABEN

*Gesetzliche Einschränkungen verhindern die Nutzung vollständiger oder aussagekräftiger Daten.*

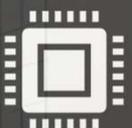
*Ein Krankenhaus darf personenbezogene Patientendaten nicht für die KI-Entwicklung nutzen, sodass wichtige Merkmale fehlen oder verfälscht werden müssen.*



# INTRANSPARENZ

*Ein Modell erzeugt Ausgaben, ohne dass nachvollziehbar ist, welche Faktoren beigetragen haben.*

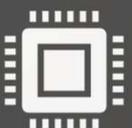
*Ein KI-Modell stuft Personen in einem Fachgebiet als Einsteiger, Fortgeschrittene oder Profis ein, ohne offenzulegen, welche Merkmale oder Kriterien zur Einstufung geführt haben.*



# FEHLKALIBRIERUNG

*Ein Modell gibt Wahrscheinlichkeiten aus, die nicht mit der tatsächlichen Unsicherheit oder den realen Verhältnissen übereinstimmen.*

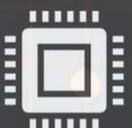
*Ein Modell zur Kreditvergabe schätzt die Rückzahlungswahrscheinlichkeit einer Person auf 90 %, obwohl sie tatsächlich nur 30 % beträgt.*



# FAIRLERNEN

*Beim Training eines Modells wird nicht überprüft, ob bestimmte Gruppen benachteiligt werden. Es fehlen Fairness-Metriken zur Kontrolle.*

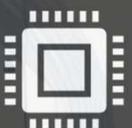
*Ein Modell zur Mieterauswahl bevorzugt unbewusst eine bestimmte Altersgruppe, da Fairness bei der Modellentwicklung nicht gemessen oder optimiert wurde.*



# RECHENKRÜCKE

*Ein Modell ist nicht dafür geeignet, die Vielfalt in den Daten angemessen zu repräsentieren, wodurch Verzerrungen entstehen können.*

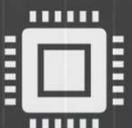
*Ein Modell zur Bilderzeugung kann Buchstaben nicht korrekt generieren, weil die Architektur nicht darauf ausgelegt ist, präzise Formen und Symbole darzustellen.*



# MANIPULATION

*Ein Modell produziert fehlerhafte Ausgaben aufgrund bewusster minimale Änderungen an Inputs, die für Menschen oft nicht erkennbar sind.*

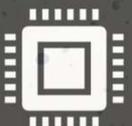
*Ein Stoppschild mit einer kleinen Verfälschung, zum Beispiel einem Aufkleber, wird von einem Modell für autonomes Fahren fälschlicherweise als „Vorfahrt gewähren“ klassifiziert.*



# UNSICHERHEIT

*Ein Modell produziert Ergebnisse ohne ausreichende Sicherheit, was zu falschen, riskanten oder schwer nachvollziehbaren Entscheidungen führen kann.*

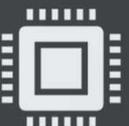
*Ein Modell für medizinische Diagnostik gibt eine Diagnose für eine seltene Erkrankung aus, obwohl es kaum passende Daten zur Erkennung dieser hat.*



# INSTABILITÄT

*Ein Modell reagiert empfindlich auf kleine Veränderungen in den Eingaben. Diese führen zu völlig anderen Ergebnissen, obwohl sich die eigentliche Situation kaum verändert hat.*

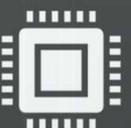
*Ein Modell zur Gesichtserkennung erkennt eine Person nicht mehr, nur weil sich die Frisur geändert hat oder die Beleuchtung anders ist.*



# MODELLSTARRE

*Ein Modell bleibt statisch und berücksichtigt keine neuen Regeln, Trends oder gesellschaftlichen Veränderungen.*

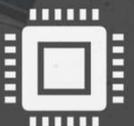
*Ein KI-Modell für Empfehlungen im Einzelhandel bevorzugt bisherige Produkte, obwohl die Nachfrage nach nachhaltigen und regionalen Produkten gestiegen ist.*



# SELBSTLÄUFER

*Ein Modell wird im laufenden Betrieb mit neuen Daten ohne menschliche Kontrolle oder definierte Grenzen trainiert und angepasst.*

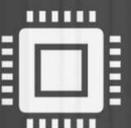
*Ein Chatbot, der aus Nutzerinteraktionen lernt, übernimmt nach einiger Zeit beleidigende oder unangemessene Sprache, weil niemand überwacht, was gelernt wird.*



# LERNRESISTENZ

*Ein Modell speichert keine neuen Erkenntnisse und wiederholt immer die gleichen Fehler.*

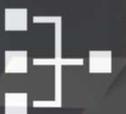
*Ein Sprachassistent versteht einen seltenen Namen immer falsch, selbst wenn Nutzer ihn mehrfach korrigieren.*



# KONTROLLVERLUST

*Ein KI-System produziert Ausgaben,  
ohne dass Menschen sie überprüfen  
oder eingreifen können.*

*Ein KI-gestütztes Buchungssystem passt  
Zimmerpreise wöchentlich automatisch  
an, ohne dass das Hotelpersonal  
eingreifen kann*



# ERKLÄRUNGSLOCH



*Ein KI-System gibt ein Ergebnis aus, ohne dass nachvollziehbar ist, welche Faktoren dazu geführt haben.*

*Ein KI-System zur Bewertung der Bonität für Wohnungsfinanzierungen lehnt einen Kreditantrag ab, ohne offenzulegen, welche Faktoren zur Entscheidung geführt haben.*



# PHANTOMSCHULD

*Es ist nicht festgelegt, wer für die Ausgaben eines KI-Systems verantwortlich ist oder Korrekturen vornehmen muss.*

*Eine Bildgenerierungs-KI erstellt irreführende oder urheberrechtlich problematische Bilder, doch niemand weiß, wer für diese Inhalte verantwortlich ist.*



# AUDITMANGEL

*Es gibt keine festgelegten Verfahren, um ein KI-System nach Einsatzbeginn unabhängig zu prüfen und nachzuvollziehen, ob es problematische Ausgaben produziert.*

*Ein KI-gestütztes System zur automatisierten Vertragsprüfung wird eingesetzt, aber es gibt keine Mechanismen, um zu überprüfen, ob es diskriminierende Tendenzen entwickelt (hat).*



# TESTLÜCKE

*Es existieren keine etablierten Prüfverfahren und Metriken, um ein KI-System während der Entwicklung zu testen.*

*Ein KI-gestütztes Job-Vermittlungssystem empfiehlt jungen Frauen vor allem Kosmetik- und Friseurberufe, weil nicht getestet wurde, ob das Modell geschlechtsneutral arbeitet.*



# DATENSCHUTZLÜCKE

*Ein KI-System verarbeitet personenbezogene Daten ohne ausreichende Schutzmaßnahmen.*

*Ein KI-gestütztes CRM-System speichert Kaufhistorien und persönliche Daten, doch es ist unklar, wer Zugriff darauf hat und wie lange diese gespeichert werden.*



# SICHERHEITSLÜCKE

*Ein KI-System ist anfällig für Manipulationen oder Angriffe, da keine ausreichenden Schutzmaßnahmen vorhanden sind.*

*Ein KI-System für Werbung im Online-Handel wird gezielt mit Fake-Klicks beeinflusst, sodass Produkte häufiger beworben werden, die nicht die beste Wahl für Kund\*innen sind.*



# WERTEVAKUUM

*Ein KI-System arbeitet ohne festgelegte ethische Richtlinien oder Grenzen.*

*Ein KI-gestütztes Empfehlungssystem verstärkt polarisierende Inhalte, weil keine ethischen Vorgaben definiert wurden, um ausgewogene Informationen zu fördern.*

# WISSENSBARRIERE

*Nutzer\*innen wissen nicht, wie sie ein KI-System richtig verwenden oder dessen Ausgaben einschätzen sollen.*

*Ein automatisierter Chatbot gibt medizinische Ratschläge, doch Nutzerinnen verstehen nicht, dass diese keine ärztliche Beratung ersetzen.*



# MISSBRAUCH

*Ein KI-System wird für unethische oder unbeabsichtigte Zwecke genutzt.*

*Eine Bildgenerierungs-KI wird genutzt, um täuschend echte Fake-Bilder von Personen ohne deren Zustimmung zu erstellen.*



# GEWINN VOR ETHIK



*Bewusster Einsatz von wirtschaftlich vorteilhaften aber ethisch fragwürdigen KI-Systemen.*

*Ein KI-gestütztes Reservierungssystem berechnet flexibel höhere Stornogebühren für gefragte Tage, selbst wenn das Zimmer noch leicht neu belegt werden könnte.*

# EINSEITIGKEIT

*Einseitige Perspektiven in der Organisation führen dazu, dass ethische Risiken unvollständig erkannt oder ignoriert werden.*

*Ein Entwicklerteam aus jungen Männern berücksichtigt bei einem KI-gestützten Gesundheitsassistenten nicht die Bedürfnisse älterer Menschen oder Frauen.*

# GRAUZONE

*Ohne klare gesetzliche Vorgaben gibt es keine einheitliche Basis für den Einsatz von KI-Systemen.*

*Ein KI-System entwirft Gebäudepläne, doch ohne gesetzliche Vorgaben ist unklar, wem die Urheberrechte gehören und wer für Fehler haftet: Die Architekt\*in oder der Softwarehersteller.*

# ETHIKDEFIZIT

*Ethische Prinzipien werden im Entwicklungsprozess nicht systematisch berücksichtigt.*

*Ein LLM-basierter Chatbot wird ohne inhaltliche Einschränkungen entwickelt, sodass er auf Fragen zu betrügerischen oder unmoralischen Handlungen bereitwillig Empfehlungen gibt.*

# EFFIZIENZDENKEN

*Unternehmen verzichten auf Prüfungen und Kontrollen, weil sie keinen direkten wirtschaftlichen Nutzen davon erwarten.*

*Ein KI-System zur Rekrutierung von Mitarbeiter\*innen wird nicht auf Diskriminierung getestet, da dies zusätzliche Kosten verursacht und gesetzlich nicht zwingend vorgeschrieben sind.*

# VORURTEILSERBE

*KI-Systeme übernehmen unkritisch die unbewussten Vorurteile und Annahmen ihrer Entwickler\*innen.*

*Ein KI-gestütztes Ernährungsberatungssystem empfiehlt hauptsächlich westliche Lebensmittel, weil das Entwicklerteam aus einer Region mit dieser Ernährungskultur stammt.*

# TUNNELBLICK



*Designentscheidungen werden getroffen, ohne die Auswirkungen auf unterschiedliche interne und externe Stakeholder berücksichtigen.*

*Ein KI-System für dynamische Preisgestaltung setzt diese rein auf Basis der Kaufhistorie fest, ohne Stammkund\*innen mit Sonderkonditionen zu berücksichtigen.*

# RISIKOIMPORT

*Ein KI-System nutzt externe Module oder Modelle, ohne dass geprüft wurde, ob diese all technischen und nicht-technischen Anforderungen erfüllen.*

*Eine Kfz-Werkstatt nutzt ein externes KI-System zur Schadensbewertung, ohne zu prüfen, ob Kratzer auf allen Lackfarben gleich zuverlässig erkannt und korrekt eingestuft werden.*

# KONTROLLDEFIZIT

*Ein KI-System automatisiert Prozesse, bei denen menschliche Kontrolle notwendig wäre.*

*Ein KI-gestütztes Versicherungssystem lehnt Anträge automatisch ab, ohne dass ein Mensch überprüft, ob es Sonderfälle oder Fehler in den Daten gibt.*

# HAFTUNGSLOCH

*Unternehmen oder Entwickler\*innen werden nicht zur Rechenschaft gezogen, wenn KI-Systeme Schaden verursachen.*

*Ein KI-gestütztes Gutachertool für Immobilienbewertungen stuft Objekte systematisch falsch ein, doch es gibt keine klare Regelung, wer für falsche Bewertungen haftet.*