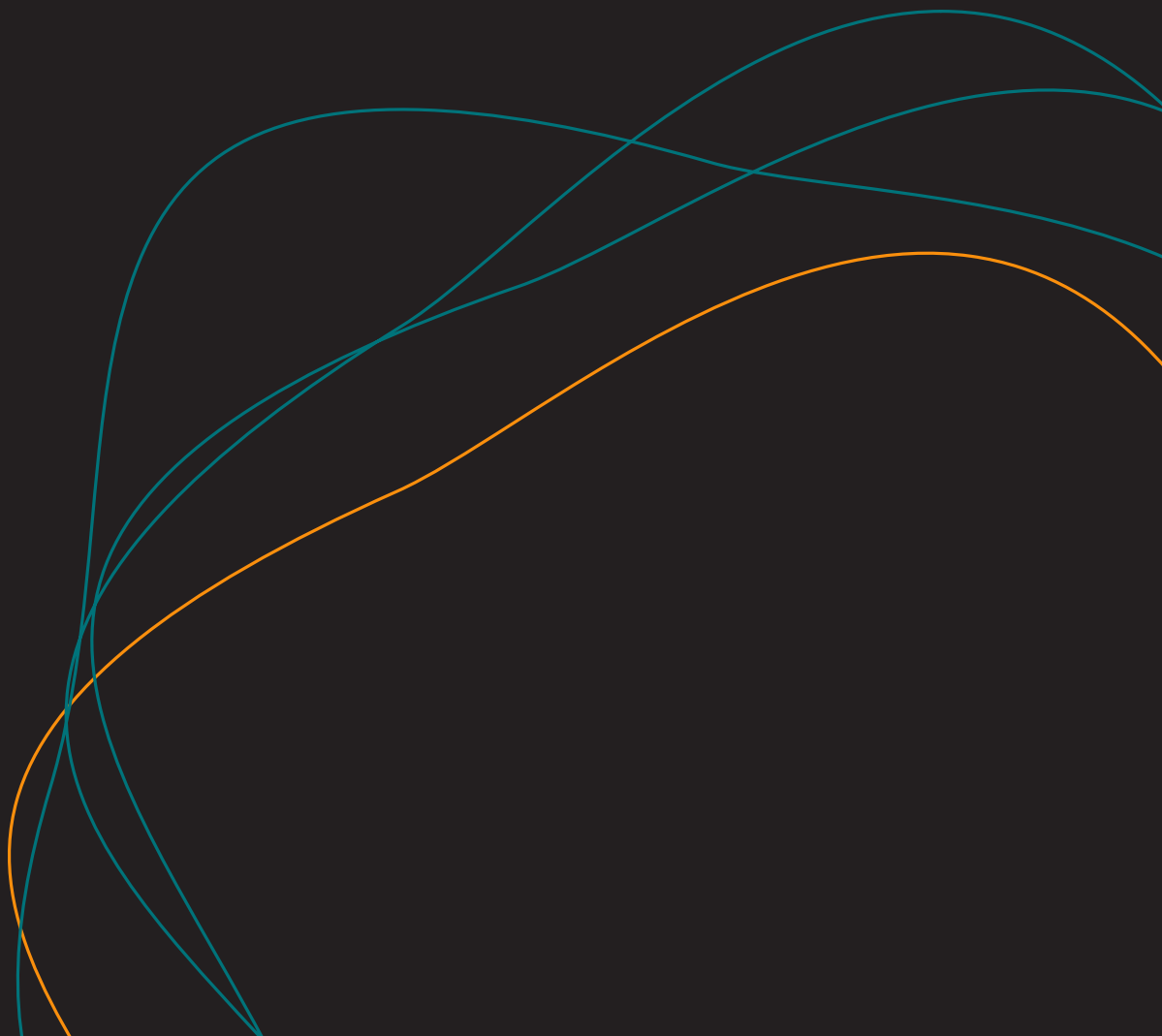
initiative for applied artificial intelligence

AI Act: Klassifizierung von KI-Anwendungen aus der Praxisperspektive

Eine Studie zur Identifikation von Unklarheiten bei KI-Anwendern auf der Grundlage von über 100 klassifizierten Systemen in Unternehmen



Inhalte

Executive Summary

Künstliche Intelligenz wird zunehmend ein Teil unseres Alltags, ob zu Hause, in der Industrie oder im öffentlichen Bereich. Die Technologie birgt Risiken, eröffnet aber auch neue Chancen. Dies stellt die Institutionen in Brüssel vor die Herausforderung, eine Balance zwischen Innovation und Regulierung für KI in der EU zu finden.

Die kommende KI-Verordnung (engl. AI Act) konzentriert sich auf die präventive Vermeidung von Schäden für Gesundheit, Sicherheit und grundlegende Menschenrechte. Konkret verfolgt sie einen risikobasierten Ansatz, wonach KI-Systeme einer Risikoklasse zugeordnet werden und Hochrisiko-Systeme strengere Anforderungen erfüllen müssen als KI-Systeme in einer niedrigen Risikoklasse.

Basierend auf dem Entwurf der EU-Kommission der KI-Verordnung (von April 2021) untersucht diese Studie, welchen Einfluss die Regeln für die Risikoklassifizierung der KI-Verordnung auf KI-Innovationen in Unternehmen haben und welche Fragen geklärt werden müssen, um für mehr Klarheit und Planungssicherheit zu sorgen. Zum Zeitpunkt der Publikation dieser Studie (März 2023) sind die Verhandlungen in Brüssel noch nicht abgeschlossen, und wir hoffen, dass unsere Vorschläge für eine präzise Klassifizierung von den verhandlungsführenden Personen aufgenommen werden (siehe nächste Seite).

Die Studie im Überblick:

- Risikoklassifizierung von über 100 KI-Systemen aus unterschiedlichen Unternehmensbereichen wie Marketing, Produktion, Einkauf etc. gemäß dem Kommissionsentwurf der KI-Verordnung von 2021 und dem Diskussionsstand im Parlament von Anfang 2022.
- 19 % der KI-Systeme zählen zur Hochrisiko-Klasse, 41 % stellen ein niedriges Risiko dar und bei 39 % ist es unklar, ob sie in die Hochrisiko-Klasse fallen oder nicht. Damit liegt der Anteil von Hochrisiko-Systemen in dieser Probe zwischen 19 % bis 58 %. Eines der KI-Systeme ist möglicherweise verboten.
- Die meisten Hochrisiko-Systeme befinden sich in den Unternehmensbereichen Personalwesen, Kundenservice, Buchhaltung und Finanzen sowie Rechtswesen, was aus Kostengründen und erhöhter Komplexität zu einer weiteren Adaptionsbarriere führen kann. Daher profitieren in diesen Bereichen tendenziell weniger Unternehmen von KI.
- Unklare Risikoklassifizierungen bremsen Investitionen und Innovationen. Die Ursachen für unklare Risikoklassifizierungen liegen vor allem in den Bereichen Kritische Infrastruktur, Beschäftigung, Strafverfolgung und Produktsicherheit (Anhang II).
- Die Untersuchung der Ursachen für Unklarheiten resultiert in konkreten Empfehlungen an die Politik und an Unternehmen, um verantwortliche KI-Innovation zu fördern.

Hauptursachen für unklare Risikoklassifizierungen

Kritische Infrastruktur	• Es ist unklar, ob die europäischen oder nationalen Definitionen zur Bestimmung von Kritischer Infrastruktur gelten. • Es ist unklar, welche Anlagenarten und Schwellenwerte für die Bestimmung von Kritischer Infrastruktur anwendbar sind, bzw. ob diese für die skalierbare Natur von KI geeignet sind. • Es ist unklar, was als „Sicherheitskomponenten“ im Bereich Kritische Infrastruktur anzusehen ist, etwa in dezentralen Systemen wie Stromnetzen oder Bahnsystemen.
Beschäftigung	• Es ist unklar, wie der Begriff „Aufgabe“ bei Aufgabenverteilung („task allocation“) definiert ist und wie er sich z. B. von Empfehlungen abgrenzt. • Es ist unklar, wie vertragliche Arbeitsverhältnisse ausgestaltet sein müssen, damit ein KI-System in diesem Kontext als „Hochrisiko“ eingestuft bzw. nicht als solches eingestuft wird.
Strafverfolgung	• Es ist unklar, unter welchen Umständen ein KI-System „im Auftrag von einer Behörde“ zur Strafverfolgung eingesetzt wird, insbesondere wenn Unternehmen zu bestimmten Maßnahmen gesetzlich verpflichtet sind, etwa in den Bereichen Geldwäsche, Betrugserkennung, Steuererklärungen. • Es ist unklar, welche Definition von „Straftat” („criminal offence“) gilt, die nationale oder eine europäische. • Es ist unklar, unter welchen Bedingungen ein Dokument (oder eine andere Information) als Beweis oder Fakt einzustufen ist. Ist es zum Beispiel erforderlich, dass ein Verfahren bereits läuft oder sind auch Informationen gemeint, die zum Beweis werden können?
Sicherheits- komponente (insb. Anhang II)	• Es ist unklar, welche Systemgrenze bei der Feststellung, ob ein KI-System als Sicherheitskomponente („safety component“) eingesetzt ist, gilt? Bei vorausschauender Wartung ist das KI-System häufig nicht Teil von dem zu wartenden Produkt. • Es ist unklar, welche Definition von „Sicherheitskomponente“ mit Blick auf sektorspezifische Standards (z. B. Automotive, Medizinprodukte), Gesetze (z. B. BSI-Gesetz §2 (13) oder Richtlinien (z. B. 2014/33/EU über Aufzüge und Sicherheitsbauteile für Aufzüge, Anhang 3) anzuwenden ist. • Es ist unklar, ob ein KI-System für eine sicherheitskritische Funktion keine Sicherheitskomponente ist, wenn es dafür redundante Maßnahmen gibt, die bei einem Ausfall oder Fehler des KI-Systems „einspringen“ und einen Schaden verhindern.

Vorwort

Hauptursachen für unklare Risikoklassifizierungen

Sehr geehrte Damen und Herren,

Künstliche Intelligenz hat in den letzten Jahren enorme Fortschritte gemacht und sich zu einer der bedeutendsten Technologien entwickelt. Von Chatbots über maschinelles Lernen bis hin zu autonomen Systemen – KI führt zu disruptiven Entwicklungen und hat bereits heute signifikanten Einfluss auf unser tägliches Leben. Für mich ist deshalb eines von besonderer Bedeutung: Wir als Gesellschaft müssen jederzeit in der Lage sein zu entscheiden, ob, wo und wie wir KI einsetzen. KI darf keine Black Box sein.

Deshalb begrüße ich das Ziel des AI Acts, Europa zum Zentrum für vertrauenswürdige KI zu machen. Dabei müssen die berechtigten Interessen nach Innovationsförderung und dem Schutz des Einzelnen in einen angemessenen Ausgleich gebracht werden. Wir brauchen also eine Regulierung, die schützt und gleichzeitig ausreichend Raum für Innovation lässt. Ansonsten drohen wir nicht nur den Anschluss zu verlieren und uns in eine technologische Abhängigkeit von China und den USA zu begeben. Wir würden auch die Möglichkeit aufgeben, selbstbestimmt unser freiheitlich-demokratisches Wertesystem in die digitale Welt zu tragen.

Der AI Act muss zum Innovationsmotor für Europa werden. Wir brauchen endlich faire Marktzugangschancen, damit unsere europäischen KMUs und Start-ups auf Augenhöhe mit den Entwicklungen in anderen Teilen der Welt konkurrieren können.

Die vorliegende, bislang einzigartige Studie beleuchtet erstmals anhand konkreter Fälle von Unternehmen, welche Auswirkungen die Regelungen des AI Acts in der Praxis haben. Die Ergebnisse sind leider teils alarmierend und zeigen deutlich die Schwächen des Verordnungsentwurfs auf. Die Klassifizierung für den risikobasierten Ansatz bleibt beispielsweise zu unklar und erhöht dadurch unnötig die Aufwände bei den Unternehmen. Die Studie macht nicht nur auf die Schwachstellen des Verordnungsentwurfs aufmerksam, sondern zeigt auch konkrete Änderungsmöglichkeiten auf.

Noch ist Zeit zum Gegensteuern! Noch kann der AI Act tatsächlich zu einem Wettbewerbsvorteil für Europa werden. Dafür setze ich mich ein.

Judith Gerlach, MdL
Bayerische Staatsministerin für Digitales

appliedAI Initiative

Die appliedAI Initiative wurde 2017 ins Leben gerufen mit dem Ziel, die Anwendung von KI zu beschleunigen und damit Europas Industrie im KI-Zeitalter wettbewerbsfähig zu halten. Dabei sind wir der Überzeugung, dass wir bei der wichtigsten disruptiven Technologie unserer Zeit auf europäische Werte und qualitativ hochwertige KI-Systeme setzen wollen und müssen.

In diesem Kontext ist die KI-Regulierung der EU der wichtigste legislative Eingriff zur Erreichung unseres Ziels. Einerseits wird dadurch ein wertebasierter „vertrauenswürdiger" Einsatz von KI eingefordert, andererseits bedeutet dies im globalen Kontext zuerst einmal Mehraufwand und Komplexität für die europäische Industrie und damit ein systembedingter Nachteil im globalen Rennen um die KI-Führungspositionen.

Um dieser Ambivalenz gerecht zu werden, haben wir uns vorgenommen, in einem konstruktiven Dialog auf die praktischen Schwierigkeiten, die der AI Act mit sich bringt, hinzuweisen und Lösungswege aufzuzeigen. Zeitgleich arbeiten wir mit unseren Partnerunternehmen daran, möglichst umfassend und zielgerichtet die europäische Industrie mit Werkzeugen, Hilfestellungen und Expertise zu unterstützen. Diese Studie ist gemeinsam mit der vor wenigen Monaten veröffentlichten Umfrage zum Einfluss des AI Acts auf das europäische Innovationsökosystem der Beginn einer Serie an Aktivitäten zu diesem wichtigen und richtungsweisenden Vorhaben.

Erstmalig werden Anwendungsfälle von KI in den maßgeblichen internen Funktionsbereichen von Unternehmen in ihrer Risikoklasse analysiert. Die Anwendungsfälle stellen dabei den Großteil von möglichen Einsatzgebieten in Unternehmensprozessen dar und geben daher einen guten Überblick über die Auswirkung des AI Acts für Millionen von Unternehmen. Nicht analysiert wurden die oft sehr spezifischen Einsatzmöglichkeiten in Produkten oder industriespezifischen Prozessschritten.

Wir sind fest davon überzeugt, dass wir es gemeinsam schaffen können, aus dem AI Act eine Erfolgsgeschichte zu machen. Dabei müssen wir aber unser Ziel fest im Blick haben: Um gestalten zu können, müssen wir technologisch und wirtschaftlich handlungsfähig bleiben. Unser Innovationsökosystem ist unsere Zukunft. Dabei ist die Berücksichtigung von Chancen genauso wichtig wie die von Risiken.

Andreas Liebl
Managing Director of appliedAI

Motivation

Die KI-Verordnung kommt und das europäische KI-Ökosystem bereitet sich darauf vor. Die appliedAI Initiative verfolgt das Ziel, die Verhandlungen in Brüssel und die darauffolgende Umsetzung mit einer praktischen Perspektive zu informieren und arbeitet zu diesem Zweck mit anderen europäischen und nationalen Partnern zusammen.

Unsere Aktivitäten orientieren sich an folgenden Zielen:

- Unterstützung von vertrauenswürdiger KI zum Schutz von Gesellschaft und Wirtschaft
- Steigerung der Wettbewerbsfähigkeit von „AI made in Europe“
- Beschleunigung der KI-Entwicklung in Europa im Vergleich mit den USA und China

Der risikobasierte Ansatz ist ein zentraler Mechanismus der KI-Verordnung mit wesentlichen Auswirkungen auf den Einsatz von KI, weil nur Hochrisiko-KI-Systeme die grundlegenden Anforderungen erfüllen müssen. Für Anbieter und Nutzer von Hochrisiko-Systemen steigt die Komplexität bei der Entwicklung und Nutzung von KI, und damit auch die Kosten, was sich auf die Adaption von KI in der Praxis auswirkt. Die Klassifizierungsregeln beeinflussen daher einerseits die Vertrauenswürdigkeit der verfügbaren KI-Systeme sowie andererseits die Fähigkeit von Firmen, solche KI-Systeme nachhaltig zu entwickeln und zu nutzen.

Vor diesem Hintergrund verfolgt diese Studie einen explorativen Ansatz, um die vorgeschlagenen Klassifizierungsregeln aus praktischer Sicht kritisch zu prüfen.

Leitfragen:

- Welche KI-Systeme lassen sich eindeutig klassifizieren und welche nicht?
- Welche Formulierungen in den Klassifizierungsregeln führen zu Unklarheiten?
- Welche Maßnahmen verringern die rechtliche Ungewissheit und beschleunigen die Umsetzung der neuen Anforderungen?

Studiendesign

Daten

Die Datengrundlage dieser Studie sind über 100 KI-Systeme aus einer öffentlichen Bibliothek, die appliedAI im Rahmen eines BMWK-geförderten Projektes erstellt hat. Die KI-Anwendungen sind nach Unternehmensbereichen wie Marketing, Produktion oder Personalwesen gruppiert und anhand folgender Merkmale beschrieben:

- Allgemeine Beschreibung der Situation
- Das Geschäftsproblem, das mit KI adressiert werden soll
- Beschreibung des KI-Systems, inklusive Funktionsweise und Nutzerszenario
- Links und Referenzen mit Hintergrundinformationen

Hier der direkte Link zur offenen und kostenfreien Datenbank (nach Anmeldung):

https://www.appliedai.de/de/ki-kompetenz-kurs

Unternehmensfunktion	Anzahl der KI-Systeme in dieser Studie
Buchhaltung und Finanzen	10
Einkauf	8
Forschung und Entwicklung	9
IT und Sicherheit	11
Kundenservice	14
Logistik und Lieferketten	11
Marketing und Vertrieb	14
Personalwesen	10
Produktion und Herstellung	9
Rechtswesen	10
Gesamt	106

Methode

Jedes der über 100 KI-Systeme wurde gemäß dem initialen Entwurf der KI-Verordnung der EU-Kommission klassifiziert. Die Klassifizierungsregeln wurden von appliedAI und mehreren Partnerunternehmen im Rahmen einer Arbeitsgruppe zum Thema „KI-Regulierung & Governance“ operationalisiert und in eine Methode überführt.

Wichtig: Die Methode basiert auf den Klassifizierungsregeln aus dem initialen Vorschlag der EU Kommission (April 2021) und den Änderungsvorschlägen vom April 2022.

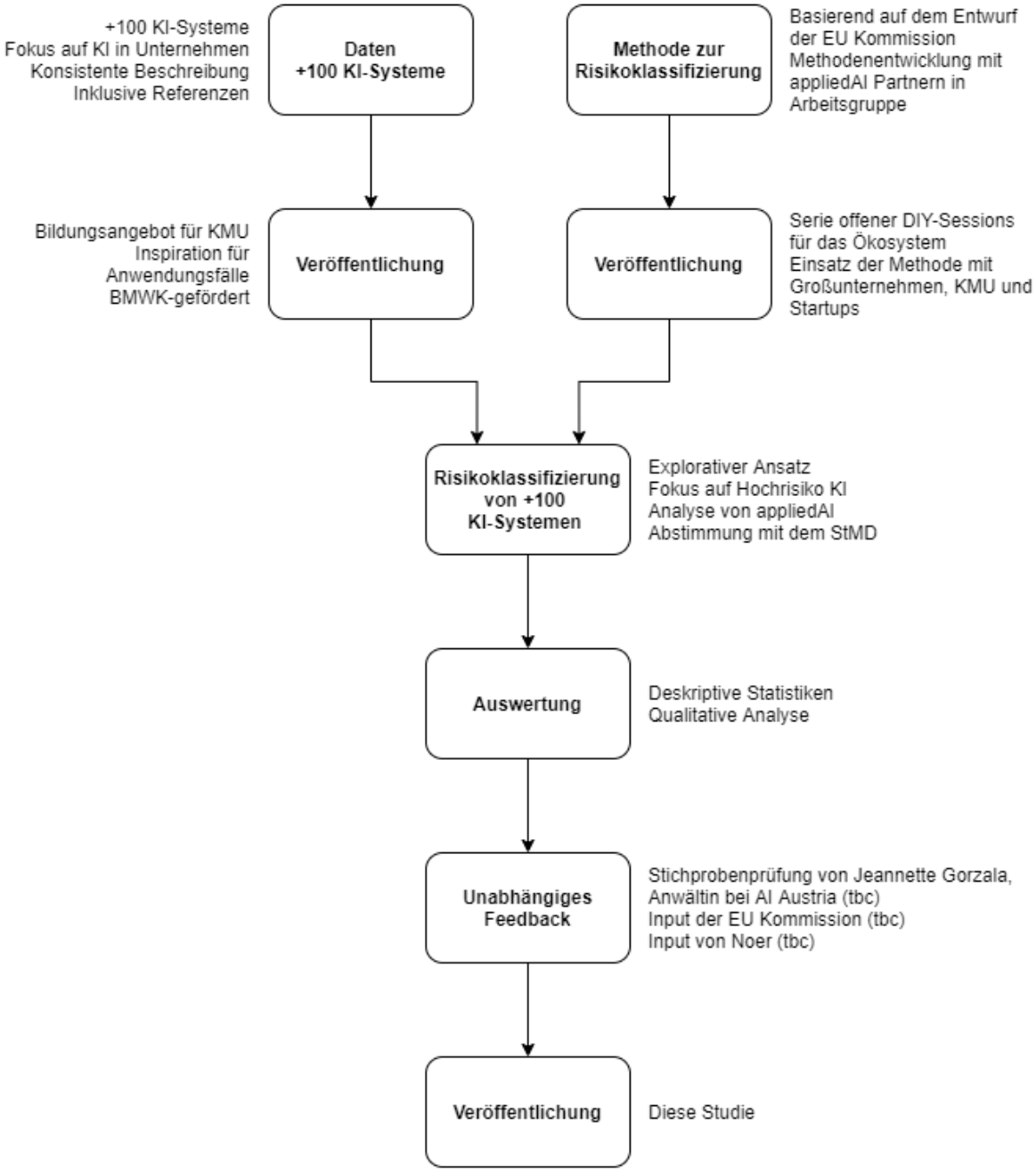
Die Methode besteht im Kern aus vier Fragen, die pro KI-Anwendung zu beantworten sind:

1. Ist das System ein KI-System im Sinne der KI-Verordnung?
2. Wenn ja, fällt das KI-System in den Anwendungsbereich der KI-Verordnung?
3. Wenn ja, ist das KI-System in der EU verboten?
4. Wenn nein, fällt das KI-System in die Hochrisiko-Klasse?

Hier der direkte Link zur Vorlage der Methode (ohne Anmeldung):

https://miro.com/app/board/uXjVOz16ydQ=/

Studienablauf



Risikoklassifizierung von über 100 KI-Systemen

Die Risikopyramide und die Annahme der 5–15 %

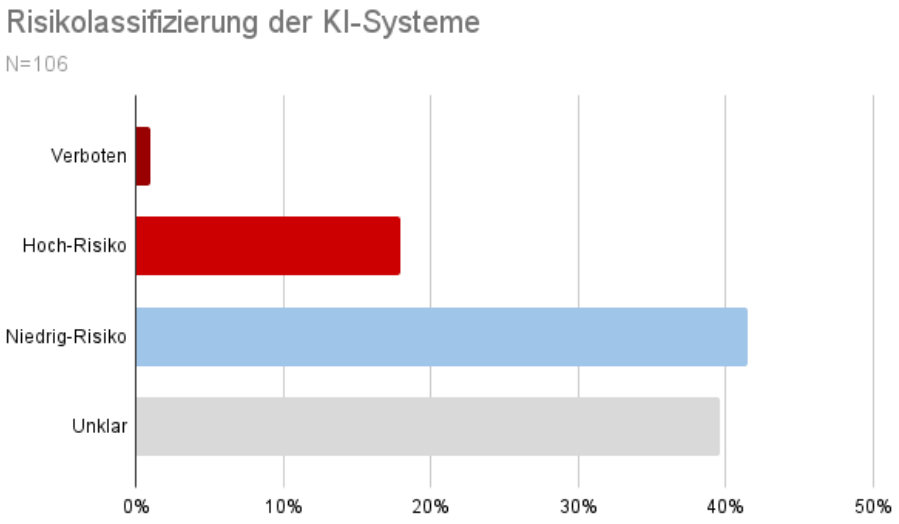
Die KI-Verordnung verfolgt einen risikobasierten Ansatz, um den Eingriff des Gesetzgebers proportional zum Risiko eines KI-Systems zu halten. Grundsätzlich sieht die KI-Verordnung drei Risikoklassen vor:

- Verbotene KI-Systeme (Artikel 5)
- Hochrisiko-KI-Systeme (Artikel 6)
- Niedrigrisiko-KI-Systeme, die nicht unter Artikel 5 oder Artikel 6 fallen

Verbotene KI-Systeme dürfen in der EU nicht genutzt oder verfügbar gemacht werden (siehe Titel 2 der KI-Verordnung). Das Inverkehrbringen von Hochrisiko-KI-Systemen ist nur möglich, wenn die Anforderungen in Titel 3 (insb. Kapitel 2) erfüllt sind. Niedrigrisiko-KI-Systeme sind nicht reguliert bzw. wird für diese ein freiwilliger Verhaltenskodex empfohlen.

Ferner sieht Artikel 52 Transparenzanforderungen für KI-Systeme vor, die mit natürlichen Personen interagieren, jedoch gilt diese Anforderung für KI-Systeme der Hoch- und Niedrigrisiko-Klasse gleichermaßen, weshalb sie hier nicht separat aufgeführt werden.

Die Risikoklassifizierung von 106 KI-Systemen aus der öffentlich zugänglichen Anwendungsfall-Bibliothek[1] von appliedAI ergibt folgende Verteilung:



Der Anteil der Hochrisiko-KI-Systeme liegt mit 18 % über dem von der EU-Kommission

1 https://www.appliedai.de/de/ki-kompetenz-kurs

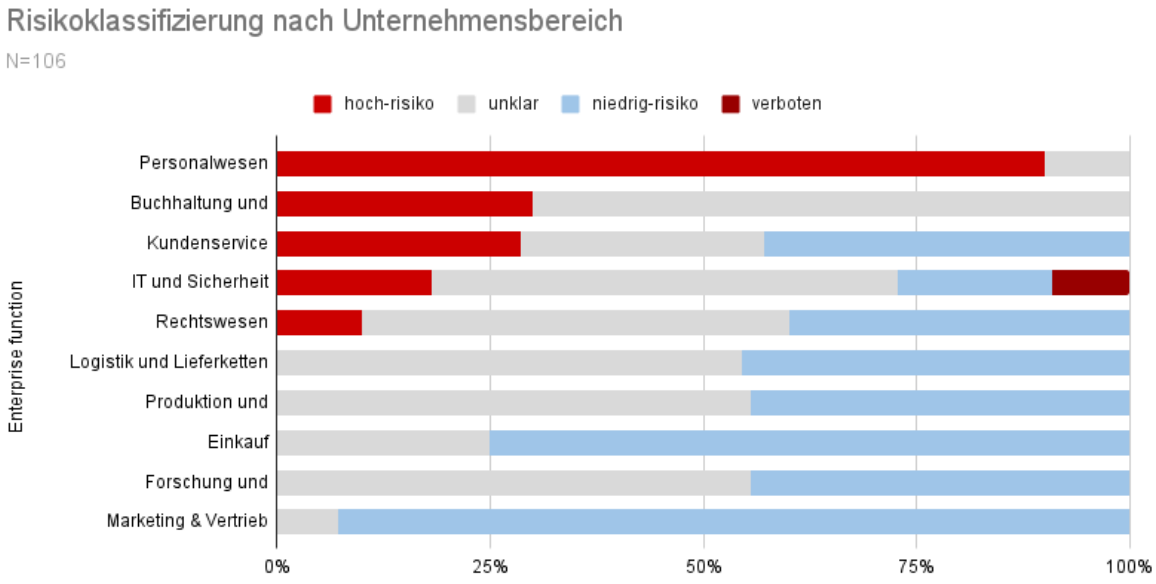
Risikoklasse	Anzahl/Count	Anteil/Share
Verboten	1	1 %
Hochrisiko	19	18 %
Niedrigrisiko	44	42 %
Unklar	42	40 %

angenommenen Maximalwert von 15 % (geschätzt wurden 5–15 %[2]) . Durch die weiteren 40 % an unklaren Fällen, bei denen ein KI-System nicht eindeutig als Hoch- oder Niedrigrisiko-KI-System klassifiziert werden konnte, liegt der potenzielle Anteil von Hochrisiko-KI-Systemen bei bis zu 58 %. Diese Entwicklung ist von zentraler Bedeutung für die Folgenabschätzung, denn die meisten Anforderungen der KI-Verordnung gelten für Hochrisiko-KI-Systeme und deren Anbieter, für die die wirtschaftlichen Gesamtkosten und -aufwände entsprechend steigen.

Auswirkung der Klassifizierungsregeln auf KI in Unternehmen

Die untersuchten KI-Systeme in dieser Studie finden in allgemeinen Unternehmensfunktionen, zum Beispiel Marketing, Produktion, Finanzen oder Personalwesen Anwendung. Solche KI-Systeme sind branchenunabhängig und nicht nur relevant für Großunternehmen, sondern auch für kleine und mittelständische Betriebe. Daher haben sie ein besonders großes Potenzial, einen Mehrwert zu schaffen. Die großflächige Nutzung von KI-Technologie kann zu spürbaren Produktivitätssteigerungen beitragen und damit bei gleichem bzw. sinkendem Ressourceneinsatz die Leistung einer Wirtschaft erhöhen. Vereinfacht gesagt, „der Kuchen wird größer“.

Abbildung XX zeigt die Verteilung der Risikoklassifizierungen nach Unternehmensbereichen.



2 Impact Assessment, Accompanying the Proposal for a Regulation of the European Parliament and of the Council LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS, EU Commission, 2021

Unternehmensfunktion	Hochrisiko	Unklar	Niedrigrisiko	Verboten	Summe
Buchhaltung und Finanzen	3	7			10
Einkauf		2	6		8
Forschung und Entwicklung		5	4		9
IT und Sicherheit	2	6	2	1	11
Kundenservice	4	4	6		14
Logistik und Lieferketten		6	5		11
Marketing und Vertrieb		1	13		14
Personalwesen	9	1			10
Produktion und Herstellung		5	4		9
Rechtswesen	1	5	4		10
Summe	**19**	**42**	**44**	**1**	**106**

Im Personalwesen sind mehr als 75 % der KI-Systeme in der Hochrisiko-Klasse eingestuft und in den Bereichen Kundenservice, Buchhaltung und Finanzen sowie IT und Sicherheit sind es jeweils mehr als 25 %. Unklare Klassifizierungen finden sich in allen Unternehmensbereichen, jedoch am meisten in Buchhaltung und Finanzen mit über 70 %. Nur im Marketing und Vertrieb liegt der Anteil der unklaren Fälle unter 25 %.

Die **Risikoklasse** eines KI-Systems hat einen Einfluss auf die Wahrscheinlichkeit, dass es entwickelt und adaptiert, bzw. finanziert wird. Hochrisiko-Systeme haben eine geringere Chance, umgesetzt zu werden, weil der Anstieg an Kosten und Komplexität durch die zusätzlichen Anforderungen die Adaptionsbarriere erhöht. Eine paneuropäische Umfrage[3] unter 113 KI-Start-ups und 15 Wagniskapitalfirmen hat gezeigt, dass viele Firmen die neuen Anforderungen zu Data Governance und Risk-Management als „schwierig" bis hin zu „sehr schwierig" einschätzen. Ferner stellt die Durchführung eines Konformitätsbewertungsverfahrens eine Herausforderung für viele Startups dar.

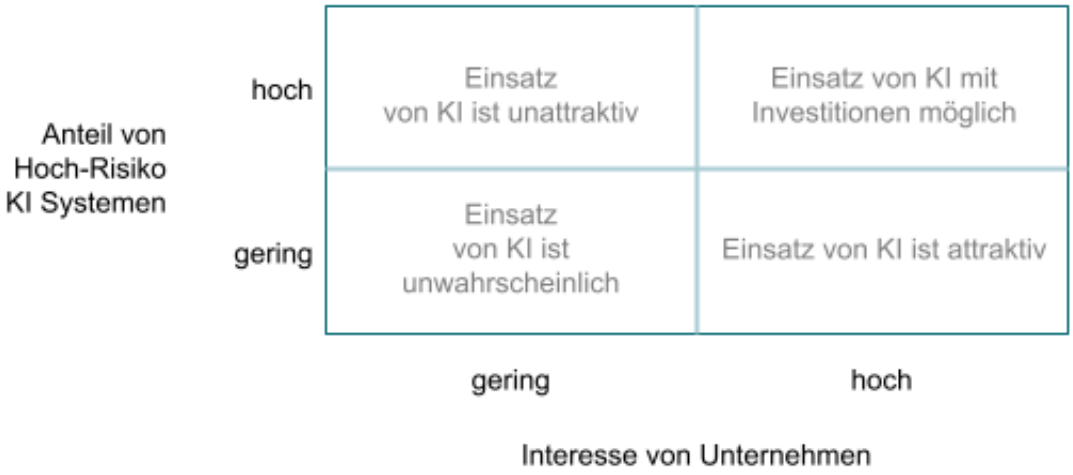
Um die wachsenden Kosten zu rechtfertigen, muss das Wertversprechen (engl. Value Proposition) eines KI-Systems entsprechend hoch sein, damit sich die Investition in dessen Entwicklung lohnt. Daher hat auch der **Mehrwert** des KI-Systems einen Einfluss auf dessen Adaption in Unternehmen. Eine Umfrage von Bitkom e.V. [4]zeigt, in welchen Unternehmensbereichen KI mit hoher Wahrscheinlichkeit eingesetzt wird. Insgesamt 539 Unternehmen beantworteten dabei die Frage „In welchen Bereichen Ihres Unternehmens kommen KI-Tools zum Einsatz bzw. in welchem Bereich halten Sie einen künftigen Einsatz für wahrscheinlich?"

Unter den Firmen, die zum Zeitpunkt der Befragung noch keine KI einsetzten, sind Bereiche wie Kundenservice („zur Kundenbindung" mit 86 %) und IT („in der IT-Abteilung" mit 82 %) besonders beliebt. Weniger interessant ist KI in der Rechts- und Steuerabteilung sowie Forschung und Entwicklung.

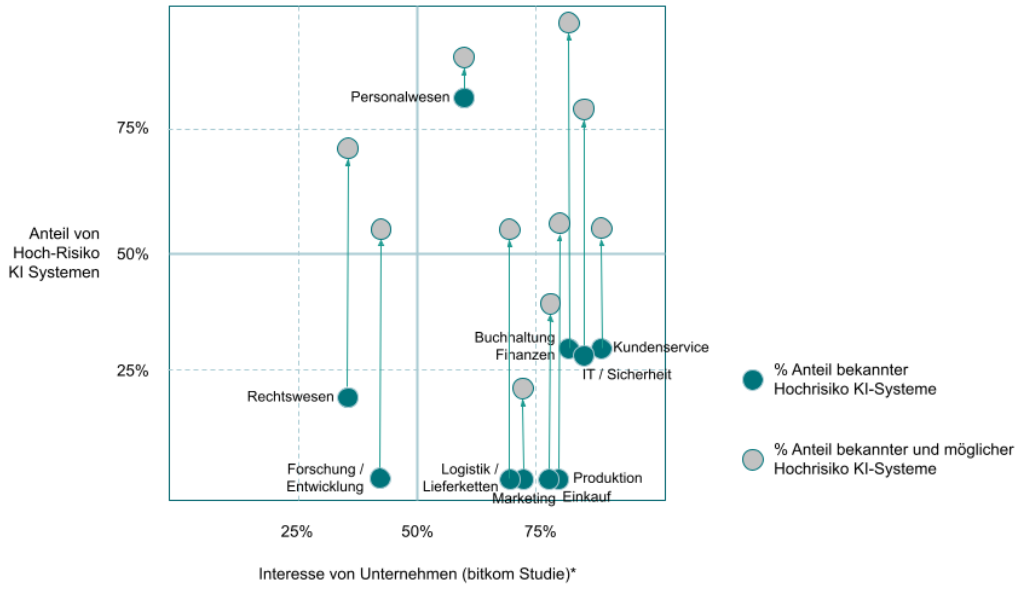
3 AI Act Impact Survey, appliedAI Initiative, 2022, https://www.appliedai.de/hub/ai-act-impact-survey

4 Künstliche Intelligenz – Wo steht die deutsche Wirtschaft?, Bitkom, 2022, https://www.bitkom.org/sites/main/files/2022-09/Charts_Kuenstliche_Intelligenz_130922.pdf

Vor diesem Hintergrund ist zu erwarten, dass KI-Systeme insbesondere in Unternehmensbereichen mit einem geringen Anteil an Hochrisiko-Systemen und einem großen Potenzial, einen Mehrwert zu stiften, zum Einsatz kommen. Im Gegensatz dazu wird KI seltener in Unternehmensbereichen mit einem hohen Risikopotenzial und einem geringen Mehrwert zu finden sein. Insgesamt lassen sich für KI in Unternehmensbereichen folgende Tendenzen ableiten (wobei diese von anderen organisatorischen Faktoren wie Expertise, Kultur, Infrastruktur oder Risikobereitschaft im Management beeinflusst werden):



Auf dieser Grundlage wurden nun die Unternehmensbereiche (Punkte) den vier Quadranten der Matrix zugeordnet, siehe Abbildung XX. Der Anteil der Hochrisiko-Systeme (y-Achse) stammt aus den Analysen von appliedAI und das Interesse von Unternehmen bezieht sich auf die Antworten der genannten der Bitkom-Studie (x-Achse). Der grüne Punkt zeigt für jeden Unternehmensbereich die Position der bestätigten[5] Hochrisiko-Systeme (Low-risk scenario) und der graue Punkt zeigt die Position für den Fall, dass alle unklaren Fälle auch als Hochrisiko-Systeme klassifiziert werden (High-risk scenario).



Hinweis: Die Grenze der Quadranten bei 50 % ist hypothetisch und sie könnte auch woanders liegen. Eine konkrete Bewertung der KI-Systeme pro Unternehmensfunktion (z. B. monetär) könnte zu einer anderen Aufteilung führen.

5 Bestätigt im Sinne, dass die Klassifizierung als Hochrisiko-System wahrscheinlich erscheint.

Das **Low-Risk-Szenario** beschreibt den Fall, dass alle unklaren Fälle in die Niedrigrisiko-Klasse fallen. Hier befinden sich mehrere Unternehmensfunktionen im Quadranten rechts unten, bei denen der Einsatz von KI attraktiv ist, da es sich mehrheitlich um Niedrigrisiko-Systeme mit einem hohen Interesse von Unternehmen handelt. Das heißt, dass Investitionen in die Entwicklung sowie Initiativen für die Anwendung wahrscheinlich sind, wodurch Unternehmen direkt von den zu erwartenden Produktivitätssteigerungen profitieren können. Dies betrifft Unternehmensfunktionen wie Logistik und Lieferketten, Marketing und Vertrieb, Produktion und Einkauf.

Das **High-Risk-Szenario** beschreibt den Fall, dass alle unklaren Fälle in die Hochrisiko-Klasse fallen. Hier verschiebt sich das Bild durch den hohen Anteil an unklaren Klassifizierungen deutlich, wodurch die meisten Unternehmensbereiche im Quadranten rechts oben landen, was wahrscheinlich zu einer Hürde für Investitionen und Adaptionen führt. Nur die Bereiche Marketing und Einkauf wären noch im „attraktiven Quadranten" rechts unten. Diese Entwicklung würde den Einsatz von KI in Unternehmensbereichen wie Buchhaltung und Finanzen, Kundenservice, IT und Sicherheit oder Produktion ausbremsen, weil sich hier der Einsatz von Hochrisiko-KI durch die gestiegenen Anforderungen womöglich nicht lohnt. In diesem Fall würden Unternehmen von den Potenzialen von KI nicht oder kaum profitieren.

Die **Auswertung** zeigt, dass der große Anteil unklarer Risikoklassifizierungen viel Ungewissheit in allen Bereichen erzeugt, die die Investitionen in KI und die ohnehin schon träge Adaption von KI in Deutschland weiter ausbremsen kann. Hier schwingt auch eine Angst vor Fehlern bzw. Strafen in den Unternehmen mit: Der zitierten Bitkom-Studie zufolge sind „Verstöße gegen Datenschutzvorgaben" die zweithäufigste Sorge beim Einsatz von KI unter den befragten Unternehmen (N=606). Diese Sorge ist bei der KI-Verordnung möglicherweise größer als bei der DSGVO, weil die KI-Verordnung neu und daher unbekannter ist und die Strafen bei Verstößen höher sind. Außerdem nennen 49 % der befragten Unternehmen die „Verunsicherung durch rechtliche Hürden" als ein Hemmnis für den Einsatz von KI.

Das Ziel dieser Auswertung ist nicht, dass KI-Systeme mit einem hohen Risikopotenzial unreguliert in Umlauf gebracht werden. Das Ziel ist es, eine vorsichtige Prognose abzugeben, welchen Einfluss die Klassifizierungsregeln auf die Adaption von KI in Unternehmen haben und wie wichtig dabei eine präzise Formulierung in der KI-Verordnung ist.

Um die Potenziale von KI für Unternehmen zu nutzen, ohne den Schutz von Gesundheit, Sicherheit und grundlegender Menschenrechte zu kompromittieren, ist es wichtig, dass die Klassifizierungsregeln in der KI-Verordnung eindeutig sind, um Unklarheiten zu reduzieren und Planungssicherheit bei Investitionen zu erzeugen. Daher folgt nun eine explorative Analyse mit dem Ziel, die Ursachen für Unklarheiten zu identifizieren, um entsprechende Gegenmaßnahmen abzuleiten.

Praktische Prüfung der Klassifizierungsregeln

In diesem Kapitel wird aus praktischer Sicht geprüft, inwiefern mit den vorgeschlagenen Klassifizierungsregeln eine eindeutige Zuordnung in die Risikoklassen möglich ist.

Für jedes der 106 KI-Systeme haben wir geprüft:

- Handelt es sich um ein Hochrisiko-System oder nicht?
- Welche Artikel, Anhänge oder Erwägungsgründe (engl. Recitals) sind zutreffend?
- Falls eine eindeutige Klassifizierung nicht möglich ist, was sind die Ursachen?

Die Klassifizierung erfolgte auf Basis folgender Vorschriften der KI-Verordnung, die wir vereinfacht zusammengefasst haben:

Verbotene KI-Systeme	Artikel 5 Erwägungsgründe 7–24 Leitfragen: 1. Setzt das KI-System unterschwellige Techniken außerhalb des Bewusstseins einer Person ein, um das Verhalten einer Person wesentlich zu beeinflussen? 2. Nutzt das KI-System die potenziellen Verwundbarkeiten einer bestimmten Gruppe aus, um deren Verhalten wesentlich zu beeinflussen? a. Wenn entweder 1 oder 2 mit „Ja“ beantwortet wird: Ist es wahrscheinlich, dass die Verhaltensänderung zu einem physischen oder psychischen Schaden bei dieser Person oder einer anderen Person führt? 3. Bewertet oder klassifiziert das KI-System die Vertrauenswürdigkeit von natürlichen Personen über einen bestimmten Zeitraum, z. B. auf der Grundlage ihres Sozialverhaltens oder bekannter oder vorhergesagter persönlicher oder charakterlicher Merkmale und führt dies zu einer ungünstigen Behandlung, die in keinem Zusammenhang mit dem Kontext steht oder ungerechtfertigt ist oder in keinem Verhältnis zu dem Verhalten oder seiner Schwere steht? 4. Setzt das KI-System biometrische Fernerkennungssysteme in öffentlich zugänglichen Räumen zum Zweck der Strafverfolgung in „Echtzeit“ ein?

Hochrisiko-KI-Systeme	Artikel 6 Erwägungsgründe 30–40 Anhang II und III Ein KI-System fällt gemäß Artikel 6 in die Hochrisiko-Klasse, wenn (zusammengefasst): • Das KI-System ein Produkt ist und in den Anwendungsbereich einer der Regularien in **Anhang 2** fällt und ein Konformitätsbewertungsverfahren durchlaufen muss, oder eine Sicherheitskomponente von einem Produkt in Anhang 2 ist, das ein Konformitätsbewertungsverfahren durchlaufen muss. • Der Verwendungszweck des KI-Systems in einen der Anwendungsbereiche von **Anhang 3** fällt.
Niedrigrisiko-KI-Systeme	Artikel 69 In die Klasse der Niedrigrisiko-KI-Systeme fallen alle KI-Systeme, die nicht nach Artikel 5 verboten und keine Hochrisiko-KI-Systeme nach Artikel 6 sind. Die Bestimmung, was ein Niedrigrisiko-KI-System ist, erfolgt nicht nach einer Definition oder anhand von Kriterien, sondern nach einem Ausschlussverfahren.

Eindeutige Klassifizierung

Dieses Kapitel zeigt KI-Systeme, die gemäß der Klassifizierungsregeln des initialen Entwurfs der KI-Kommission vom 21. April 2021 eindeutig bzw. mit einer hohen Wahrscheinlichkeit einer Risikoklasse zugeordnet werden können.

Die Auflistung dient der Illustration und soll zum Reflektieren anregen, ob die Klassifizierung und die damit einhergehenden Anforderungen korrekt bzw. angemessen sind.

Zusätzlich bildet dieses Kapitel einen Referenzpunkt für die unklaren Klassifizierungen im darauffolgenden Kapitel.

Verbotene KI-Systeme

In der Probe von 106 KI-Anwendungen haben wir ein (potenziell) verbotenes KI-System identifiziert:

ID 93

Unternehmensbereich: IT und Sicherheit

Name: Erkennung von Bedrohungen bei Großveranstaltungen

Kontext:

Flughäfen bleiben ein prototypisches Ziel für Angriffe mit Massenopfern. Viele Flughäfen im ganzen Land und weltweit haben große Investitionen in Technologie, Personal und Prozesse getätigt, um Reisende und andere Flughafengäste zu überprüfen. Als zweitgrößter Flughafen in Nordkalifornien ist der Passagierverkehr am Oakland International Airport (OAK) auf dem besten Weg, die 13,2 Millionen Reisenden zu übertreffen, die im vergangenen Jahr den Flughafen passierten. Um diesem Wachstum gerecht zu werden, stellte der OAK zusätzliches Personal ein, um den Flughafen sicher zu halten, und begann mit der Erforschung innovativer Lösungen in Bezug auf Methoden und Ausrüstung für die Inspektion von Mitarbeitenden. Der OAK begann, nach einer neuen Ausrüstungsplattform zu suchen, die in der Lage ist, eine größere Auswahl an potenziellen Waffen zu erkennen und gleichzeitig die betriebliche Effizienz mit zunehmender Anzahl der Mitarbeitenden zu verbessern. Der Status quo – eine Kombination aus zeitaufwendigen, manchmal invasiven Maßnahmen, einschließlich begehbarer Metalldetektoren und gelegentlichen Ganzkörperabtastungen – würde wahrscheinlich zu langen Warteschlangen zu Beginn jeder Schicht und daraus resultierend zu sinkender Arbeitsmoral führen.

KI-System:

Die KI-Dienste eines externen Partners, die mit einem umfangreichen Satz realer Bedrohungsdaten trainiert wurden, unterscheiden kontinuierlich echte Bedrohungen von harmlosen Objekten in Echtzeit. Sie werden mit der Zeit intelligenter, wenn neue Bedrohungsprofile entdeckt werden. Darüber hinaus zeigen sie dem Sicherheitspersonal genau, wo Waffen am Körper der Person oder in ihrer Tasche versteckt sein könnten, und ermöglichen es den Wachen, präzise und schnell einzugreifen. Die Technologie verwendet künstliche Intelligenz und Gesichtserkennungssoftware, um Live-Aufnahmen von sich nähernden Flughafengästen zu analysieren und so festzustellen, ob es sich um zugelassene Personen handelt, wie z. B. regelmäßig wiederkehrende Flughafengäste, VIPs, Mitarbeitende und andere Personen, denen Zutritt gewährt werden sollte. Wenn eine nicht zulässige Person von Interesse hervorgehoben wird, wird ihr Profil an Sicherheitsbeauftragte gesendet und eine menschliche Fachkraft kann die Daten überprüfen und verifizieren. Die Technologie beansprucht, mindestens einer Person pro Sekunde den Zutritt zu ermöglichen.

Begründung:

Das KI-System dient zur Erkennung von Personen und Gegenständen mit Gefährdungspotenzial an einem öffentlich zugänglichen Raum. Es verwendet biometrische Erkennung und gibt Informationen an Sicherheitspersonal, u. a. um „nicht zugelassene“ Personen manuell zu überprüfen. Die Ausnahmen zum Verbot in Artikel 5 Abs.1 lit d) ii und iii treffen möglicherweise zu*, aber dazu liegen nicht ausreichend Informationen über das KI-System vor.

*

ii) Verhinderung von Gefahren für Kritische Infrastruktur, Gesundheit, Sicherheit und Leben
iii) Identifikation von Personen, die im Rahmen eines Verfahrens gesucht werden

Verbotene KI-Systeme

In der Probe von 106 KI-Anwendungen haben wir ein (potenziell) verbotenes KI-System identifiziert:

ID 42

Unternehmensbereich: Kundenservice

Name: Intelligente Suche Beispiel 2

Anhang 3: 1. Biometric identification

Kontext:

Das Finanzinstitut hat sich 2017 mit einem externen Anbieter zusammengetan, um in seinen Contact Centern Sprachbiometrie für die Authentifizierung bereitzustellen. Nach einer reibungslosen Erfahrung, die zu einer stärkeren Personalisierung und einer schnelleren Lösung bei Live-Anrufen von Agent:innen führte, wollte das Vermögensverwaltungsunternehmen noch innovativer sein und seinen Wettbewerbsvorteil weiter ausbauen. Dies geschah durch Authentifizierung von Anrufer:innen in ihrem Interactive Voice Response System (IVR), noch bevor sie einer/m Agentin/Agenten zugestellt wurden. Ihre bestehende Sprachbiometrielösung wurde um das IVR erweitert und das System so abgestimmt, dass es Anrufer:innen anhand minimaler Sprachäußerungen authentifiziert.

Magento ist eine Open-Source-E-Commerce-Plattform. Die Standardsuche von Magento eignet sich gut für grundlegende Anwendungsfälle: Sie bietet eine automatische Vervollständigung und Synonyme können hinzugefügt werden. Ihr Ziel war es jedoch, eine bessere Benutzererfahrung zu bieten. Magento wollte die Benutzer:innen mit Strategien, wie dem Vermarkten von Suchbegriffen oder der Personalisierung von Ergebnissen, eng mit den richtigen Inhalten zur richtigen Zeit verbinden.

KI-System:

Magento hat sich mit einem externen Anbieter zusammengetan, um seiner Kundschaft eine intelligente Suche auf seiner Website zu ermöglichen. Auf der Support-Center-Seite der Website werden Besucher:innen hervorgehobene Teile der Suchergebnisse angezeigt, die ihrer Anfrage entsprechen. Durch sogenanntes Highlighting und Snipping zeigt Magento den Suchenden, warum ein bestimmtes Ergebnis verzögert wird, und erleichtert so die Suche nach der richtigen Lösung.

Begründung:

Das KI-System ist ein Hochrisiko-System nach Anhang 3 Abs. 1 lit a), weil die Identifikation anrufender Personen anhand ihrer Stimme eine Form biometrischer Erkennung ist.

ID 88

Unternehmensbereich: IT und Sicherheit

Name: Erkennung ausgeklügelter Cyber-Angriffe

Anhang 3: 2. Kritische Infrastruktur

Kontext:

Nach einer Zeit der Unternehmensumstrukturierung suchte Energy Saving Trust, eine unabhängige Organisation, die sich mit Energieeffizienz und sauberen Energielösungen beschäftigt, nach einer Cyber-Sicherheitstechnologie, um ihre gesamte Cyber-Abwehrstrategie zu stärken. Der Trust war bestrebt, seine kritischen Vermögenswerte, einschließlich sensibler Kundendaten und geistigem Eigentum, vor ausgeklügelten und intelligenten Cyber-Angriffen zu schützen, und erkannte die Notwendigkeit einer Technologie, die selbst die subtilsten Bedrohungen erkennen kann.

Darüber hinaus verwaltet der Trust ein dynamisches und komplexes Netzwerk, was ihn natürlich anfällig für potenzielle Insider-Bedrohungen macht. Der Trust erforderte eine vollständige Netzwerktransparenz, um ungewöhnliches Verhalten sofort erkennen zu können, sei es von einer ahnungslosen Arbeitskraft, deren System gehackt wurde, oder einer Person mit böswilliger Absicht, die zur Nutzung des Systems autorisiert ist.

KI-System:

Der Energy Saving Trust arbeitete mit einem externen Anbieter zusammen, um eine auf KI-Technologie basierende Plattform zu entwickeln. Die entstandene Plattform modelliert das Verhalten jedes Geräts, Benutzers und Netzwerks, um bestimmte Muster zu lernen. Das System erkennt automatisch jedes anomale Verhalten und alarmiert das Unternehmen in Echtzeit. Es tut dies, ohne sich auf voreingestellte Regeln oder Signaturen verlassen zu müssen, wie die meisten Legacy-Tools, und erkennt daher potenzielle Bedrohungen mit größerer Wahrscheinlichkeit, auch wenn sie zuvor nicht aufgetreten sind. Energy Saving Trust kann so zahlreiche anomale Aktivitäten erkennen, sobald sie auftreten, und das Sicherheitsteam alarmieren, um weitere Untersuchungen durchzuführen, während gleichzeitig jedes Risiko gemindert wird, bevor wirklicher Schaden entsteht.

Begründung:

Unter der Annahme, dass der Energy Saving Trust Aufgaben im Bereich der Stromerzeugung und verteilung übernimmt, ist das KI-System nach Anhang 3 Abs. 2 lit a) eine Hochrisiko-Anwendung, weil es in der Form einer Sicherheitskomponente zum Schutz kritischer Infrastruktur dient.

Hinweis: Nach Erwägungsgrund (34) der Kompromissversion des Rats vom 25. November 2022 wäre dieses KI-System keine Hochrisiko-Anwendung (vgl. „Components intended to be used solely for cybersecurity purposes should not qualify as safety components.“).

ID 2

Unternehmensbereich: Personalwesen

Name: Analyse eines Videointerviews

Anhang 3: 4. Arbeit und Mitarbeit

Kontext:

Das Kundenbetreuungskonzept ist ein zentrales Element des Geschäftsmodells der Mietwagen-Vergleichsplattform HAPPYCAR. Daher ist es besonders wichtig, zuverlässige und schnelle Erkenntnisse über die Persönlichkeit und Kommunikationsfähigkeit von Personen zu erhalten, die sich für diesen Funktionsbereich bewerben. HAPPYCAR suchte nach einer

Lösung, um die Bewertungsqualität zu erhöhen und den Rekrutierungsprozess zu beschleunigen.

KI-System:

HAPPYCAR hat die KI-Lösung von Retorio zur Analyse von Videointerviews in seinen Einstellungsprozess integriert. Basierend auf Computervisions- und Klassifizierungstechniken entwickelt Retorio ein einzigartiges Persönlichkeitsprofil basierend auf dem Big-5-Framework und ein separates Kommunikationsprofil für jeden Kandidaten/jede Kandidatin. Diese beiden Profile können dann zu einem Profil kombiniert werden, das einen umfassenden Überblick über die jeweilige Person bietet. Personen, die sich für eine Stelle in der Kundenbetreuung bei HAPPYCAR bewerben, müssen zusätzlich zum Lebenslauf ein 1-minütiges Video bereitstellen, in dem sie die Frage beantworten, warum sie bei HAPPYCAR arbeiten wollen. Diese Videos werden dann von der KI-Lösung von Retorio analysiert.

Begründung:

Das KI-System dient der Bewertung und Einstellung natürlicher Personen (Bewerber:innen) und ist damit nach Anhang 3, Abs. 4 lit a) eine Hochrisiko-Anwendung.

ID 113

Unternehmensbereich: Buchhaltung und Finanzen

Name: Risikobewertung

Anhang 3: 5. Zugang zu grundlegenden privaten und öffentlichen Leistungen

Kontext:

Ein großes Finanzunternehmen wollte Frühwarnsignale finden, um zu erkennen, ob seine Gläubiger:innen wahrscheinlich in Insolvenz gehen würden. Die traditionellen Überwachungssysteme, die sie verwendeten, überprüften die Gläubiger:innen, indem sie ihre Bankkonten, Überweisungen oder Jahresabschlüsse überprüften. Bei der Anwendung solcher Methoden geriet das Unternehmen jedoch bereits zum Zeitpunkt der Entdeckung der Warnzeichen in eine finanzielle Notlage. Das Finanzunternehmen arbeitete mit Deloitte Tschechien zusammen, um ein Frühwarnsystem für Kreditmigrationen zu schaffen.

KI-System:

Deloitte Tschechien hat ein KI-Tool namens Eagle Eye entwickelt, das Open-Source-Intelligenz verwendet, um Signale aus dem Internet zu sammeln. Die KI-Software betrachtet jede Information, die sie über das Unternehmen, den Kundenstamm oder den Markt findet, als Signal. Mithilfe von maschinellem Lernen beginnt Eagle Eye dann, diese Signale zu analysieren und zu korrelieren und kann so bestimmte Muster erkennen. KI ist in der Lage, mit den riesigen Datenmengen im Internet umzugehen und Korrelationen zwischen Parametern zu finden, an die Menschen nicht einmal denken würden. Sobald diese Muster bestimmt sind, überwacht Eagle Eye ständig das Internet, um nach ihnen Ausschau zu halten und Warnungen bereitzustellen.

Begründung:

Soweit zu den Gläubigern auch natürliche Personen zählen und das KI-System für die Kreditwürdigkeitsprüfung dieser natürlicher Personen verwendet werden soll, wäre das KI-System gemäß Anhang 3 Abs. 5 lit. (b als Hochrisiko-KI-System einzustufen. Die Ausnahme in Anhang 3 Abs. 5 lit. (b (bzw. Erwägungsgrund 37) hilft hier nicht, weil das KI-System nicht von einem kleinen oder

ID 107

Unternehmensbereich: Buchhaltung und Finanzen

Name: Betrugsaufdeckung Beispiel 1

Anhang 3: 6. Strafverfolgung

Kontext:

Die Danske Bank musste in den letzten Jahren mehrere Milliarden Euro an Bußgeldern zahlen, weil sie sich nicht an alle Finanzvorschriften und -regeln hielt. Obwohl Finanzkriminalität vorkommt, handelt es sich bei den meisten identifizierten Fällen nicht um Betrug, sondern um Fehlalarme, die durch veraltete IT-Systeme verursacht werden. Gleichzeitig werden einige tatsächliche Betrugsfälle nicht erkannt. Alle Verdachtsfälle müssen manuell durch Compliance-Beauftragte geprüft werden. Aus diesem Grund hat die Danske Bank die Zahl seiner Compliance-Mitarbeitenden innerhalb der letzten zwei Jahre auf 1.700 Mitarbeitende verdoppelt. Die starke Abhängigkeit von manueller Arbeit erhöht die Kosten erheblich.

KI-System:

Hawk:AI bekämpft Finanzbetrug mit einer Anti-Geldwäsche-Lösung, die auf Echtzeit-Transaktionsüberwachung basiert und maschinelles Lernen in Kombination mit klassischen regelbasierten Ansätzen anwendet. Ihr System analysiert und bewertet große Datensätze historischer und Echtzeit-Transaktionen. Basierend auf den Erkenntnissen aus historischen Verdachtsfällen ist das KI-System in der Lage, relevante Fälle in Echtzeit zu filtern und zur weiteren Untersuchung durch menschliche Compliance-Beauftragte zu kennzeichnen. Darüber hinaus integriert Hawk:AI neue Methoden zur automatischen Mustererkennung, die die Entdeckung neuer und unbekannter Betrugsarten ermöglichen.

Begründung:

Das KI-System hat den Zweck zu bewerten, ob eine Straftat (z. B. Geldwäsche, Betrug) begangen wurde und ist damit ein Hochrisiko-System nach Anhang 3 Abs. 6 lit a).

Niedrigrisiko-KI-Systeme

ID 26

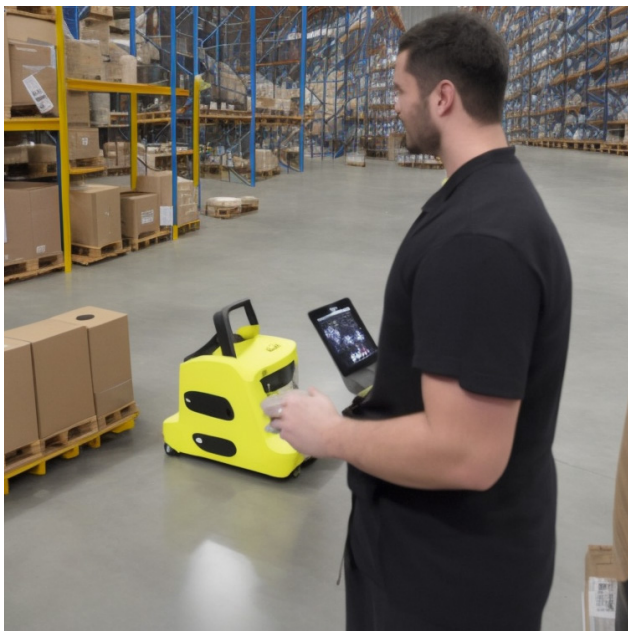
Unternehmensbereich: Einkauf

Name: Intelligente Lieferantenauswahl und -verwaltung

Kontext:

Heidelberger hatte in der Vergangenheit ein seltenes Gussteil aus einer Hand bezogen. Angesichts von Lieferengpässen und der Suche nach Möglichkeiten zur Kostensenkung, musste das Unternehmen alternative Lieferfirmen finden, denen es vertrauen konnte, und war offen für neue Lieferfirmen

aus dem Ausland. Der Auswahlprozess der richtigen Lieferfirmen, die zuverlässig sind und zeitnah Qualitätsprodukte zum richtigen Preis liefern, kann eine mühsame Aufgabe sein und auch eine Menge Analysen und Hintergrundprüfungen erfordern. Eine zusätzliche Herausforderung war die Auswahl einer Lieferfirma aus dem Ausland. Deshalb entschied sich Heidelberger für die Supplier-Intelligence-Lösung von Scoutbee.



KI-System:

Die Market- und Supplier-Intelligence-Lösung von Scoutbee bietet Einblicke in neue Märkte und erhöht mit ihren Kernprodukten die Transparenz und Informationsqualität. Die Lösung kann eine große Auswahl an potenziellen Lieferfirmen zur Verfügung stellen, gezielte Listen kuratieren und einen KI-gestützten Lieferantenvertrauenswert entwickeln. Basierend auf diesem Score und den Profilen der Lieferfirmen von Scoutbee können Unternehmen potenzielle Lieferfirmen schneller qualifizieren und bewerten.

ID 98

Unternehmensbereich: Forschung und Entwicklung

Name: Überprüfung von Geschäftsmodellen

Kontext:

Ein Hersteller von Premiumfahrzeugen wird derzeit durch Entwicklungen in den Bereichen Elektromobilität, autonomes Fahren und Carsharing herausgefordert. Das Know-how in der Verbrennungstechnik, das seit Jahrzehnten ein gewisses Alleinstellungsmerkmal garantierte, verliert schnell an Bedeutung. Sollten sich in den kommenden Jahren autonomes Fahren und Carsharing als Trend etablieren, könnte dies auch zu deutlich geringeren Fahrzeugverkäufen führen. Der Hersteller prüfte daher neue Geschäftsideen der eigenen Mitarbeitenden, der Kundschaft und externer Start-ups auf Machbarkeit und Potenzial. Die Herausforderung bestand darin, das Potenzial von Geschäftsideen richtig einzuschätzen und zu priorisieren.

KI-System:

Der Hersteller schickte sechs Mitarbeitende zu einem Workshop bei einem externen Unternehmen, das eine Lösung zur Analyse der Geschäftsideen bietet. Nach der Einführung in das Funktionsprinzip der Lösung wurden alle Mitarbeitenden gebeten, ihre Geschäftsidee nach ihrem Bauchgefühl zu bewerten. Anschließend wurden zwei Gruppen mit je drei Personen gebildet. Jeder Teilnehmende zog sich zunächst mithilfe des KI-Tools zu einer Selbsteinschätzung der Geschäftsidee zurück. Anschließend trafen sich die beiden Teams erneut und besprachen die Ergebnisse. Das KI-Tool bot konkrete Ansatzpunkte, um die Ideen zu verbessern. Darüber hinaus lernte das Team an nur einem Tag die Entscheidungsparameter erfolgreicher Venture-Capital-Investoren kennen und konnte diese zur Bewertung und Verbesserung von Geschäftsideen nutzen.

ID 26

Unternehmensbereich: IT und Sicherheit

Name: Überwachung der Datenqualität

Kontext:

Große Datensätze haben große Datenqualitätsprobleme. Für Personen, die mit Millionen von Datenpunkten zu tun haben, ist es eine Herausforderung zu wissen, wo eine Änderung stattfindet, da so viele Permutationen, Geschäftsmetriken und Dimensionen existieren. Diese Datensätze müssen auf einer Analyseplattform verarbeitet werden, die Erkennungsalgorithmen in mehreren Schritten der Datenpipeline effizient ausführen kann, um Datenqualitätsprobleme und Änderungen der Geschäftsmetriken zu identifizieren.

KI-System:

Die groß angelegte KI-Analyselösung in Echtzeit von Anodot ist in jedem Schritt des Datenerfassungsprozesses vollständig automatisiert (Erkennung, Rangfolge und Gruppierung) und gibt präzise Warnungen über Änderungen in wichtigen Geschäftskennzahlen wie fehlende Daten, unerwartete Datentypen, Nullen, wo keine sein sollten, oder fehlerhafte Aufzeichnungen. Besteht anhand dieser Benachrichtigungen der Verdacht, dass mit den Daten nicht alles in Ordnung ist, kann sich die verantwortliche Person schnell direkt auf das konkrete Problem konzentrieren und überlegen, wie sie weiter vorgeht. Diese Mehrgliedrigkeit kann Unternehmen dabei helfen, sehr spezifische Anomalien in der Datenqualität zu erkennen, insbesondere solche, die durch breitere Metriken wie Durchschnittswerte und unternehmensweite Gesamtwerte geglättet oder unbemerkt bleiben würden.

ID 49

Unternehmensbereich: Kundenservice

Name: Ursachenanalyse

Kontext:

Der Kundenservice wird oft mit Servicetickets überschwemmt. Wenn Mitarbeitende jedoch jede einzelne eingehende Anfrage lesen müssen, ist es oft unmöglich, alle Probleme zu beantworten, was zu einem Anstieg der unbeantworteten Tickets führt. Eine gute Möglichkeit, den Kundendienst zu verbessern, besteht darin, die Arbeitsbelastung der Mitarbeitenden bei wiederkehrenden Problemen zu reduzieren. Der Weg, die Mitarbeitenden zufriedener zu machen, besteht darin, die Ursache ihrer Probleme zu finden. Oftmals hat der Support-Dienst in den Tausenden von Support-Tickets, die er jede Woche erhält, einen Schatz an Informationen. Alles in Form von unstrukturierten Textdaten. Da einige Kund:innen dazu neigen, ihre Probleme ausführlich im Detail zu erklären, ist jede Ursachenanalyse, die auf dem Durchlesen der Tickets und Schätzungen basiert, ineffektiv und zeitaufwendig.

Ein E-Commerce-Unternehmen, das Druckerzeugnisse verkauft, hatte eine große Anzahl von Reklamationen zu verspäteter Lieferung, die Unzufriedenheit der Kund:innen stieg an. Jedes Kundenproblem musste individuell bearbeitet werden, und es war offensichtlich, dass diese Beschwerden enorme Zeit in Anspruch nahmen. Da die Mitarbeitenden damit beschäftigt waren, sich um die Beantwortung der Reklamationen zu kümmern, konnte das zugrunde liegende Problem nicht erkannt werden.

KI-System:

Mit der Lösung eines externen Anbieters nutzte das E-Commerce-Geschäft eine KI-gestützte Ursachenanalyse, um interessante Zusammenhänge und Ursachen zu finden, die ihm halfen, ein tieferes Problem unter der Oberfläche zu erkennen. Bei der Anzeige der Support-Tickets, im Zusammenhang mit Kundenbeschwerden über verspätete Lieferungen, waren diese Tickets eng mit Lieferfirmen verknüpft. Beschwerte sich eine Person aufgrund einer verspäteten Lieferung, nannte sie auch das Versandunternehmen. Die Erkenntnisse daraus können genutzt werden, um Versandunternehmen zu identifizieren, die mit unverhältnismäßig vielen Reklamationen in Verbindung gebracht werden, und entsprechende Maßnahmen zu ergreifen.

ID 26

Unternehmensbereich: Logistik und Lieferketten

Name: Ermöglichung vorausschauender Logistik

Kontext:

Otto ist ein wichtiger Player im deutschen E-Commerce-Markt. Eine Analyse seiner Daten zeigte, dass Retouren weniger wahrscheinlich sind, wenn die Ware innerhalb von zwei Tagen geliefert wird. Außerdem bevorzugt die Mehrheit der Kundschaft, ihre Bestellung auf einmal zu erhalten und nicht in mehreren Sendungen. Aber für Otto ist es nicht einfach, diese Faktoren zu bedienen, da die Firma Produkte verschiedener Marken verkauft, die sie nicht selbst auf Lager hat. Normalerweise bedeutet dies, entweder mit dem Versand zu warten, bis alle Produkte gesammelt sind, oder mehrere Kartons zu versenden, die zu unterschiedlichen Zeiten ankommen.

KI-System:

Die Lösung für diese Probleme besteht darin, besser vorherzusagen, was die Kundschaft kaufen wird, sodass diese Produkte im Voraus bestellt werden können. Um dies zu erreichen, verwendet Otto einen Deep-Learning-Algorithmus, der ursprünglich für Teilchenphysik-Experimente am CERN in Genf entwickelt wurde. Der KI-Algorithmus analysiert rund 3 Milliarden vergangene Transaktionen und 200 Variablen wie vergangene Verkäufe, Suchanfragen auf Otto.de sowie externe Informationen wie Wettervorhersagen, um vorherzusagen, was Verbraucher:innen kaufen werden. Das System kann jetzt mit sehr hoher Genauigkeit vorhersagen, was im nächsten Monat verkauft wird, und ermöglicht es, jeden Monat rund 200.000 Artikel automatisch von Drittmarken, ohne menschliches Eingreifen, zu bestellen. Bei Otto hat das KI-System zu einer deutlichen Reduzierung der Produktretouren geführt.

ID 68

Unternehmensbereich: Marketing

Name: Überwachung von Markenerwähnung / Social Listening

Kontext:

Somersby, eine führende Apfelweinmarke der dänischen Brauerei Carlsberg Group, wollte ihre Marketingkampagnen optimieren und besser verfolgen. Sie haben zahlreiche Hashtag-Kampagnen entwickelt, die Fans erfolgreich eingebunden haben, und ihren Erfolg durch den Aufbau starker Beziehungen zu Bloggern und Influencern verstärkt. Als sie beispielsweise eine neue Somersby-Sorte

auf dem polnischen Markt einführten, arbeiteten sie mit Dutzenden von Bloggern zusammen und ermutigten die Menschen, Inhalte (insbesondere Fotos) mit einem speziellen Hashtag zu teilen.

KI-System:

Somersby nutzte eine KI-gestützte Social-Listening-Lösung eines externen Anbieters, um diese Kampagne zu verfolgen und die Stimmung gegenüber der Marke zu überprüfen. Dank dieser Methode konnten sie sehen, dass die Kampagne die allgemeine Markenstimmung verbesserte und eine enorme Reichweite in den sozialen Medien erzielte. Darüber hinaus wurde das neue Getränk zu einem Bestseller in seiner Kategorie.

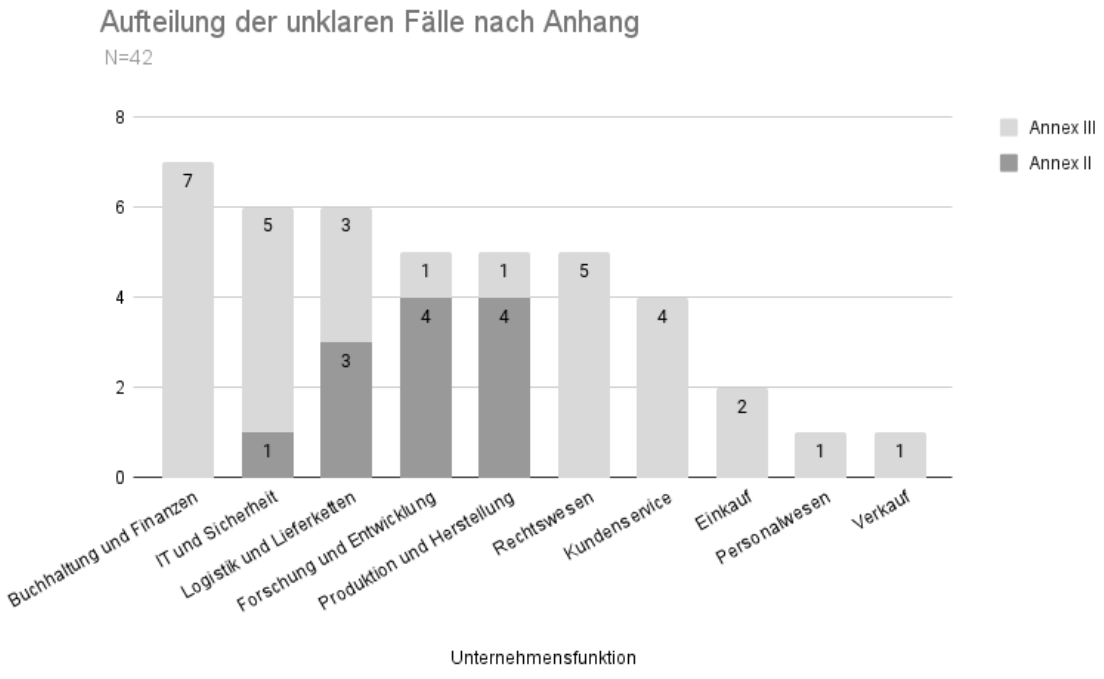
Unklare Klassifizierung

Dieses Kapitel fokussiert sich auf KI-Systeme, bei denen unklar ist, ob sie als Hochrisiko-Systeme einzustufen sind oder nicht.

Nach einer Übersicht der 42 unklaren Fälle folgt eine Untersuchung konkreter KI-Systeme in den Bereichen Kritische Infrastruktur, Beschäftigung, Strafverfolgung und Anhang 2. Pro Anwendungsbereich werden zunächst die einschlägigen Klassifizierungsregeln gelistet. Danach folgen konkrete KI-Systeme inklusive Erläuterungen, warum die Klassifizierung unklar ist. Die anschließende Diskussion greift mögliche Gründe für Unklarheiten auf und bietet eine Grundlage für die Empfehlungen für die Anpassung der Klassifizierungsregeln.

Überblick

Von den 42 KI-Systemen mit unklarer Klassifizierung fallen 30 in Anhang III und 12 in Anhang II.



Innerhalb von Anhang III gibt es die meisten Unklarheiten bei der Klassifizierung von KI-Systemen

in den Bereichen 6) Strafverfolgung, 4) Beschäftigung und 2) Kritische Infrastruktur, siehe Tabelle XX. Zusammen mit den unklaren Fällen aus Anhang II führen die Formulierungen von vier Bereichen der KI-Verordnung zu mehr als 80 % der unklaren Klassifizierungen.

Die betroffenen KI-Systeme werden im folgenden Abschnitt vorgestellt, inklusive einer Begründung für die Unklarheiten. Ziel ist es, ein genaueres Verständnis für die Ursachen der Unklarheiten zu bekommen.

Unklare Fälle nach Anhang 3:

	Sektion in Anhang 3:						
Unternehmensfunktion	**2**	**3**	**4**	**5**	**6**	**8**	**Summe**
Buchhaltung und Finanzen				2	5		7
Einkauf	1				1		2
Forschung und Entwicklung					1		1
IT und Sicherheit	2				3		5
Kundenservice			3			1	4
Logistik und Lieferketten	2		1				3
Marketing und Vertrieb		1					1
Personalwesen		1					1
Produktion und Herstellung			1				1
Summe	5	2	5	2	13	3	30

Rechtsgrundlagen der KI-Verordnung

Die Klassifizierungsregeln sind die Grundlage, um Hochrisiko-Systeme von anderen KI-Systemen zu unterscheiden, sodass unklare Formulierungen in der KI-Verordnung einen direkten Beitrag zu Verunsicherung der involvierten Akteure haben. Daher enthalten die folgenden Tabellen die einschlägigen Klassifizierungsregeln der Bereiche mit vielen unklaren Klassifizierungen, als Referenz für die darauffolgende Diskussion.

Kritische Infrastruktur
Unternehmensfunktion(en): Lieferketten und Logistik, IT und Sicherheit
Klassifizierungskriterien (Anhang III; AI Act proposal April 2021; engl.) Section 2: Critical infrastructure (+ Erwägungsgrund 34) (a) AI systems intended to be used as safety components in the management and operation of critical digital infrastructure, road traffic and the supply of water, gas, heating and electricity;

Beschäftigung

Unternehmensfunktion(en): Personalwesen, Kundenservice, Lieferketten und Logistik

Klassifizierungskriterien (Anhang III; AI Act proposal April 2021; engl.)

3. Education and vocational training (+Erwägungsgrund 35):

(a) AI systems intended to be used to determine access, admission or to assign natural persons to educational and vocational training institutions or programmes at all levels;

(b) AI systems intended to be used to evaluate learning outcomes, including when those outcomes are used to steer the learning process of natural persons in educational and vocational training institutions or programmes at all levels

4. Employment, workers management and access to self-employment (+Erwägungsgrund 35):

(a) AI systems intended to be used for recruitment or selection of natural persons, notably to place targeted job advertisements, to analyze and filter job applications, and to evaluate candidates;

(b) AI [sic] intended to be used to make decisions on promotion and termination of work-related contractual relationships, to allocate tasks based on individual behavior or personal traits or characteristics and to monitor and evaluate performance and behavior of persons in such relationships

Strafverfolgung

Unternehmensfunktion(en): Buchhaltung und Finanzen, IT und Sicherheit, Rechtswesen

Klassifizierungskriterien (Anhang III; AI Act proposal April 2021; engl.)

6. Law enforcement (+Erwägungsgrund 38):

(a) AI systems intended to be used by law enforcement authorities or on their behalf to assess the risk of a natural person for offending or reoffending or the risk for a natural person to become a potential victims of criminal offences;

(b) AI systems intended to be used by law enforcement authorities or on their behalf as polygraphs and similar tools or to detect the emotional state of a natural person;
(...)

(d) AI systems intended to be used by law enforcement authorities or on their behalf to evaluate the reliability of evidence in the course of investigation or prosecution of criminal offences;

(e) AI systems intended to be used by law enforcement authorities or on their behalf to predict the occurrence or reoccurrence of an actual or potential criminal offence based on profiling of natural persons as referred to in Article 3(4) of Directive (EU) 2016/680 or to assess personality traits and characteristics or past criminal behaviour of natural persons or groups;

(f) AI systems intended to be used by law enforcement authorities or on their behalf to profile of natural persons as referred to in Article 3(4) of Directive (EU) 2016/680 in the course of detection, investigation or prosecution of criminal offences;

8. Administration of justice and democratic processes (+ Erwägungsgrund 40):

(a) AI systems intended to be used by a judicial authority or on their behalf to interpret facts or the law to apply the law to a concrete set of facts.

Anhang 2

Unternehmensfunktion(en): Produktion, Lieferketten und Logistik, IT und Sicherheit

Klassifizierungskriterien (Anhang II; AI Act proposal April 2021; engl.)

Artikel 6 (+ Erwägungsgründe 30-31):

1. An AI system that is itself a product covered by the Union harmonisation legislation listed in Annex II shall be considered as high risk if it is required to undergo a thirdparty conformity assessment with a view to the placing on the market or putting into service of that product pursuant to the above mentioned legislation.

2. An AI system intended to be used as a safety component of a product covered by the legislation referred to in paragraph 1 shall be considered as high risk if it is required to undergo a third-party conformity assessment with a view to the placing on the market or putting into service of that product pursuant to above mentioned legislation. This provision shall apply irrespective of whether the AI system is placed on the market or put into service independently from the product.

Erwägungsgründe 27-32

Artikel 3

(14) 'safety component of a product or system' means a component of a product or of a system which fulfills a safety Unternehmensbereich for that product or system or the failure or malfunction of which endangers the health and safety of persons or property;

Kritische Infrastruktur

Logistik und Lieferketten

ID 74

Name: Flottenmanagement Beispiel 1

Kontext:

Linde ist ein weltweit tätiges multinationales Chemieunternehmen. Mehr als eine Milliarde Kilometer legen die Lieferwagen des Unternehmens jährlich zurück. Als Teil eines breiteren Fokus auf künstliche Intelligenz (KI) und Prozessoptimierung entwickelt das Unternehmen jetzt Lösungen, um die Sicherheit dieser Fahrten zu verbessern.

KI-System:

Das Sammeln von Daten ist ein Schlüsselelement jedes betrieblichen Prozesses, denn ohne die Analyse vergangener Daten lassen sich keine fundierten Entscheidungen treffen. Mit historischen Einsichten werden Millionen von in Echtzeit analysierten Datenpunkten untersucht. Daraus resultiert die Priorisierung von Chancen und Risiken, sodass Flottenmanager:innen und Fahrer:innen die beste Vorgehensweise in potenziell problematischen Situationen bestimmen können. Durch die Zusammenarbeit mit einem britischen Start-up (KI-Experten im Transportbereich) hatte Linde Zugriff auf umfangreiche Daten und begann, damit einen neuen Algorithmus zu entwickeln. Das Projekt konzentrierte sich eher auf externe Faktoren als auf Informationen zu den Fahrer:innen selbst. Linde hatte Zugriff auf die Daten des öffentlichen Verkehrs der letzten 10 Jahre, darunter zwei Millionen in Polizeiberichten beschriebene Unfälle, Straßentopologiedaten, Wetterdaten, Straßenbaudaten und Verkehrsdaten sowie Lindes eigene Fahraufzeichnungen. Durch maschinelles Lernen war es möglich, Korrelationen zwischen verschiedenen Faktoren zu identifizieren, irrelevante Informationen zu entfernen und vorherzusagen, was unter bestimmten Bedingungen am wahrscheinlichsten passieren wird.

Was ist unklar?

Die Trainingsdaten enthalten sicherheitskritische Daten (inklusive Polizeiberichte), um mit dem KI-System mögliche Gefahren für Fahrer:innen von LKWs vorherzusagen. Unklar ist, ob das KI-System als sicherheitskritische Komponente im Management des Straßenverkehrs nach Anhang 3 Abs. 2 lit a) betrachtet wird, da ein Ausfall oder eine Fehlfunktion zu einem Anstieg an Schaden führen kann.

ID 75

Name: Flottenmanagement Beispiel 2

Kontext:

Das amerikanische multinationale Technologieunternehmen Amazon nutzt viele verschiedene Transportdienste, um Pakete zu liefern. Amazon steht seit Langem in der Kritik, seine Fahrer:innen dazu zu drängen, täglich bis zu 200 Lieferungen zu erledigen, was nach Ansicht vieler eine unangemessene Forderung ist, die dazu führen kann, dass müde Fahrer:innen Risiken eingehen. Anstatt diese intensiven Zeitpläne zu reduzieren, hat das Unternehmen damit begonnen, mit KI ausgestattete Kameras zu verwenden, um Fahrer:innen zu warnen, wenn sie gegen Straßenverkehrsregeln verstoßen oder unsichere Fahrpraktiken anwenden.

KI-System:

Amazon installiert die Driveri-Plattform des in San Diego ansässigen Start-ups Netradyne in seinen Fahrzeugen. Deren Kameras verwenden vier Objektive, die die Straße, den/die Fahrer/in und beide Seiten des Lieferwagens filmen. Die Kameras, die zu 100 Prozent in Betrieb sind, zeichnen kein Audio auf und können nicht verwendet werden, um die Fahrer:innen in Echtzeit zu beobachten. Sie verfügen über künstliche Intelligenz, die 16 Signale basierend auf dem, was um das Fahrzeug herum passiert, und den Aktionen eines Fahrers/einer Fahrerin identifiziert. Alles Illegale, wie z. B. Nichtanhalten oder zu schnelles Fahren, löst Audio-Antworten aus, darunter „Kein Halt erkannt“ und „Bitte langsamer fahren“. Unsicheres Fahren, wie z. B. zu starkes Bremsen, erzeugt keine Audiowarnungen, sondern wird im Filmmaterial festgehalten. Dies wiederum wird auf ein sicheres Portal hochgeladen, damit Amazon es prüfen kann. Während die Kameras keinen Live-Feed bieten, können einige Signale Amazon dazu veranlassen, die Fahrer:innen zu kontaktieren. Wenn zum Beispiel ein Gähnen registriert wird, weist die Kamera darauf hin, für 15 Minuten anzuhalten. Wenn Fahrer:innen dies nicht tun, vermutlich wegen der Lieferungen, die abgeschlossen werden müssen, könnte die/der Vorgesetzte anrufen und sie bitten, eine Weile anzuhalten.

Hinweis: Diese Anwendung wurde erheblich kritisiert und stellt einige ethische Herausforderungen dar, da einige Fahrer:innen eine Bedrohung ihrer Privatsphäre sehen.

Was ist unklar?

Das KI-System dient u. a. dazu, Unfälle im Straßenverkehr zu vermeiden. Ein Ausfall oder Fehler des Systems kann zu einer erhöhten Gefahr für Fahrer:innen beitragen. Unklar ist, ob das KI-System nach Anhang 3 Abs. 2 lit a) als sicherheitskritische Komponente im Management des Straßenverkehrs betrachtet wird.

IT und Security

ID 87

Name: Erkennung ausgeklügelter Cyber-Angriffe

Kontext:

Vor einigen Jahren war ein Serverbruch beim Rohstoffhändler ED&F MAN Deutschland GmbH ein Weckruf für den zunehmenden Erfolg von Cyberangriffen und das damit verbundene Risiko für sensible Daten. Eine unabhängige Bewertung führte dazu, dass das Unternehmen seine Cybersicherheitsprozesse und Tools erheblich verbessern und Mitarbeitende schulen musste.

KI-System:

Das Unternehmen suchte nach einer KI-basierten Plattform eines externen Anbieters zur Erkennung und Reaktion auf Bedrohungen. Diese sammelt und speichert Netzwerk-Metadaten und reichert sie mit einzigartigen Sicherheitserkenntnissen an. Die Plattform verwendet diese Metadaten zusammen mit Techniken des maschinellen Lernens, um Angriffe in Echtzeit zu erkennen und zu priorisieren. Dies half ED&F MAN Holdings, mehrere Man-in-the-Middle-Angriffe zu erkennen und zu blockieren und ein Krypto-Mining-Programm in Asien zu stoppen. Darüber hinaus wurde Command-and-Control-Malware gefunden, die sich seit mehreren Jahren versteckt hatte.

Was ist unklar?

Das KI-System unterstützt den Schutz digitaler Infrastruktur eines Unternehmens, aber es ist unklar, ob bzw. welche Unternehmen als kritische digitale Infrastruktur im Sinne von Anhang 3 Abs. 2 der KI-Verordnung gelten. Auf der Website des BSI werden Unternehmen im besonderen öffentlichen Interesse (UBI) genannt, wozu auch die nach ihrer inländischen Wertschöpfung größten Unternehmen Deutschlands sowie wesentliche Zulieferfirmen für diese Unternehmen gehören.

ID 90

Name: Schwachstellenmanagement

Kontext:

Für den Anbieter von Business-Beratungsservices Aprio war die risikobasierte Priorisierung von Schwachstellen eine zeitaufwendige Herausforderung, da Mitarbeitende manuell bewerten mussten, was im einzigartigen Kontext jeder Umgebung wichtig war, und in der Lage sein mussten, die

Behebungsergebnisse für jede einzelne Schwachstelle zu messen. Die Herausforderung wurde durch die Komplexität verschärft, die durch Hybrid- und Multi-Cloud-Infrastrukturen entsteht, bei denen Unternehmen Schwierigkeiten haben, zu verstehen, welche Assets sich in ihrem Netzwerk befinden, was der erste Schritt im Schwachstellenmanagement ist.

KI-System:

KI ermöglicht einer Bedrohungserkennungssoftware, wie ein Hacker zu denken. Sie kann Software helfen, Schwachstellen zu identifizieren, die Cyberkriminelle normalerweise ausnutzen würden, und sie dem Nutzenden melden. Im Gegensatz zu herkömmlichen Methoden ermöglicht KI der Bedrohungserkennungssoftware außerdem, Schwachstellen in Benutzergeräten besser zu lokalisieren, bevor eine Bedrohung überhaupt aufgetreten ist. KI-gestützte Sicherheit geht über traditionelle Methoden hinaus, um besser vorherzusagen, was ein Hacker als Schwachstelle betrachten würde. In Zusammenarbeit mit einem externen Anbieter ist Aprio in der Lage, Assets automatisch zu erkennen und fortschrittliches maschinelles Lernen anzuwenden, um das Risiko zu bewerten, dass Schwachstellen im Kontext einer bestimmten Umgebung darstellen. Der bewertete Health Score verfolgt auch das Risikoprofil von Kund:innen und bietet eine kontinuierliche, greifbare Messung des Schwachstellenmanagements und der Behebungsbemühungen anstatt einer Momentaufnahme.

Was ist unklar?

Das KI-System unterstützt den Schutz digitaler Infrastruktur diverser Unternehmen nach Anhang 3 Abs. 2 lit a), indem es Schwachstellen aufdeckt. Zu den Unternehmen, die diese Leistung in Anspruch nehmen, können auch Firmen aus dem Bereich der kritischen Infrastruktur zählen. Bei gleichem Verwendungszweck (Schwachstellenerkennung) kann das KI-System in Abhängigkeit des Nutzenden (User) als Hochrisiko- oder Niedrigrisiko-Anwendung eingestuft werden. Unklar ist, ob die Klassifizierung nun dem Anbieter obliegt (z. B. durch eine Einschränkung auf ausschließlich Niedrigrisiko-Anwendungen) oder ob durch die KI-Verordnung eine Klassifizierung als Hochrisiko-Anwendungen vorgegeben ist.

Hinweis: Erwägungsgrund 34 vom Entwurf des EU-Rats würde das KI-System ggfs. als Niedrigrisiko-System klassifizieren („Components intended to be used solely for cybersecurity purposes should not qualify as safety components“).

Kritische Infrastruktur

Weiterbildung und Beförderung

ID 8

Unternehmensbereich: Personalwesen

Name: Individuelle Lernwege in der Personalentwicklung

Kontext:

Obwohl der technologische Fortschritt das L&D in Organisationen in den letzten zehn Jahren revolutioniert hat, gibt es immer noch einige übliche Probleme, denen L&D-Fachleute gegenüberstehen, wobei eine der größten Herausforderungen der Mangel an personalisiertem Lernen ist. Mitarbeitende haben begonnen, etwas anderes zu erwarten, wenn sie zur Arbeit

kommen. Sie wollen ein personalisiertes Erlebnis, kein Standarderlebnis. Sie möchten, dass Prozesse maßgeschneidert sind und für sie funktionieren.

Ein weiterer wichtiger Faktor ist, dass jede Person einen bevorzugten Lernstil hat und mit einer bestimmten Methode am effektivsten lernt. Dies kann durch Video-Tutorials, schriftliche Inhalte, persönliche Schulungen, Gamification, audiogeführte Präsentationen oder Weiteres erfolgen.

Interne Umfragen von IBM ergaben, dass mehrere Führungskräfte angesichts der schnellen digitalen Transformation Schwierigkeiten hatten, die Fähigkeiten ihres Personals auf dem neuesten Stand und relevant zu halten. Darüber hinaus entdeckten sie auch, dass es eine Ausweitung der Jobrollen gegeben hatte und, dass es notwendig war, diese multidimensionalen Jobrollen und sich verändernden Demografien am Arbeitsplatz anzugehen. Sie suchten nach einer Lösung, die als mehrdimensionale Lösung funktionieren würde, die Mitarbeitende, Interessengruppen, Inhalte, Dienste und Anbieter:innen mit einer zentralen digitalen Plattform verbindet, um zahlreiche verschiedene Rollen und Anforderungen zu erfüllen. Das traditionelle Top-down-Lernmanagement, das entscheidet, wer was wissen muss, bot den Mitarbeitenden nur begrenzte Möglichkeiten.

KI-System:

Mithilfe von KI hat IBM „Your Learning“ entwickelt, einen reichhaltigen, personalisierten digitalen Marktplatz für das Lernen. Dieser ermöglichte es den Mitarbeitenden, zu den bei ihren Teammitgliedern beliebtesten Lerninhalten zu navigieren, sich für gezielte Lernkanäle anzumelden und die Fähigkeiten und Auszeichnungen zu erkunden, die sie benötigten, um sich auf die begehrtesten Rollen des Unternehmens vorzubereiten.

Die „Your Learning“-Plattform trug auch dazu bei, den demografischen Wandel der Belegschaft anzugehen, die Lernerfahrung der Mitarbeitenden zu verbessern, die Karrieretransparenz zu fördern und die Sozialverträglichkeit in der Organisation zu verbessern. Ein lernender Chatbot steht ebenfalls rund um die Uhr zur Verfügung, um Fragen zu beantworten. Infolgedessen verzeichnete die KI-gesteuerte Lernplattform von IBM einen Anstieg der Anmeldungen und Kursabschlüsse, wodurch der strategische Erwerb von Fähigkeiten beschleunigt wurde.

KI-gestützte Analysen und Empfehlungen können genutzt werden, um die Lernbedürfnisse eines jeden Mitarbeitenden zu personalisieren. Maßgeschneiderte Kurse können auch basierend auf den Fähigkeiten, Fortschritten, Lernbedürfnissen, Qualifikationsanforderungen und bisherigen Lernerfolgen jedes einzelnen Teammitglieds entwickelt werden.

Was ist unklar?

Das KI-System erzeugt, basierend auf bestehenden Fähigkeiten und Lernerfolgen, individuelle Empfehlungen für Weiterbildungsangebote. Damit werden 1) Lernergebnisse bewertet und 2) verwendet, um den weiteren Lernprozess zu leiten (vgl. Anhang 3 Abs. 3 lit a) „AI systems intended to be used … to assign natural persons to … programmes at all levels; to steer the learning process of natural persons“), jedoch ist unklar, ob das auch auf interne/nicht offizielle Bildungsangebote zutrifft.

Ferner könnte das KI-System in die Hochrisiko-Klasse fallen, weil die Fähigkeiten der Beschäftigten mit Blick auf eine Beförderung auf eine „begehrte“ Stelle bewertet werden (vgl. Anhang 3 Abs. 3 lit b) „AI intended to be used to make decisions on promotion and termination of work-related contractual relationships …)“

Mit Erwägungsgrund 35 dürfte es entscheidend darauf ankommen, ob das jeweilige KI-System „über den Verlauf der Bildung und des Berufslebens einer Person entscheiden“ kann. Es kommt als wohl auf die Erheblichkeit und die Auswirkungen der KI-Entscheidung an, jedoch ist unklar, was hier als Maßstab gilt.

Aufgabenzuweisung („to allocate tasks“)

ID 44

Unternehmensbereich: Kundenservice

Name: Chatbot Beispiel 1

Kontext:

MAGGI ist eine internationale Marke für Gewürze, Instantsuppen und Nudeln. Sie verkauft eine Vielzahl von Produkten auf der ganzen Welt. Ihr Ziel war es, die Kundenbindung zu erhöhen. Sie stellte fest, dass es eine Frage gibt, die Kund:innen täglich interessiert: Was soll ich heute kochen? Daher wollte das MAGGI-Kochstudio seinen Kund:innen mit Rezepten und Kochtipps helfen.

KI-System:

Dazu hat MAGGI einen Chatbot namens KiM („Kitchen Intelligence by MAGGI“) entwickelt, mit dem Kund:innen über Facebook Messenger oder WhatsApp interagieren können. Kund:innen können basierend auf Zutaten, die sie zu Hause haben, Ernährungspräferenzen und -beschränkungen, Schwierigkeitsgrad und Zubereitungszeit angeben, wonach sie suchen. KiM sortiert dann aus 2500 verschiedenen Rezepten Empfehlungen aus. Außerdem kann KiM Fragen zum Kochen beantworten und zum Beispiel erklären, wie man eine Ananas am besten schält. KiM nutzt NLP und maschinelles Lernen, um logische Strukturen einer Konversation zu verstehen, Abfragen zu durchsuchen und im Laufe der Zeit zu automatisieren. Dabei lernt und speichert KiM die Präferenzen der Benutzer:innen und wird so von Dialog zu Dialog intelligenter und hilfreicher.

Was ist unklar?

Das KI-System gibt den Nutzern individuelle Empfehlungen nach ihren Präferenzen (vgl. „individual behavior or personal traits or characteristics“) und erteilt dann Anweisungen, was zu tun ist, z. B. wie man eine Ananas schält. Das KI-System scheint sich primär an Kund:innen des Gewürzherstellers zu richten und nicht an Mitarbeitende (was sich aber nicht gegenseitig ausschließt), jedoch kann mit Blick auf die Formulierung von Anhang 3 Abs. 4 lit b) das vorliegende KI-System nicht eindeutig als „nicht Hochrisiko“ bezeichnet werden, weil es sich um ein „AI Systems intended to be used to … allocate tasks based on individual behavior or personal traits or characteristics …“ handelt.

ID 34

Unternehmensbereich: Kundenservice

Name: Automatisches Anrufmanagement / Intelligent Call Routing

Kontext:

Swisscom ist ein bedeutender Telekommunikationsanbieter in der Schweiz. Sie ist die führende Anbieterin von Mobilfunk-, Netzwerk-, Internet- und Digital-TV-Diensten für Unternehmen und Privatkund:innen in der Schweiz. Wesentlich für das Unternehmen sind 4.000 Vertriebs- und Kundendienstmitarbeitende, die jährlich mehr als 50 Millionen Kontakte, meist eingehende Anrufe, bearbeiten – in Deutsch, Französisch, Italienisch und Englisch. Sie bearbeiten auch E-Mails, Chats und Briefe.

KI-System:

Durch KI hat Swisscom Kundenanrufe besser mit den leistungsstärksten Agent:innen für

verschiedene Arten von Interaktionen abgestimmt. Durch den Wechsel zwischen traditionellem und vorausschauendem Routing konnte das Unternehmen den Effekt genau messen. Die durchschnittliche Bearbeitungszeit wurde um 3 % reduziert. Darüber hinaus verwendet das Unternehmen intelligente Anrufweiterleitung, um nicht nur die durchschnittliche Bearbeitungszeit zu reduzieren, sondern auch sicherzustellen, dass Kund:innen direkt mit Agent:innen mit den richtigen Kenntnissen und Fähigkeiten verbunden werden. Es gab dabei keine negativen Auswirkungen auf andere KPIs, wie die Geschwindigkeit der Beantwortung und die Anzahl abgebrochener Anrufe.

Was ist unklar?

Das KI-System ist ein sogenanntes Decision Support System (DSS), das auf der Grundlage der Problemdiagnose Handlungsoptionen vorschlägt. Das KI-System empfiehlt die nächstbeste Vorgehensweise und leitet eingehende Anrufe an die am besten geeigneten Call-Center-Agent:innen weiter. Jedoch ist unklar, ob diese Art der Aufgabenverteilung („task allocation“) unter Anhang 3 Abs. 4 lit b) der KI-Verordnung fällt.

ID 76

Unternehmensbereich: Logistik und Lieferketten

Name: Flottenmanagement Beispiel 3

Kontext:

Linde liefert die CO2-Zylinder, die in Kneipen für Bierzapfanlagen und andere Getränkespender verwendet werden, um ihnen die sprudelnde Güte zu verleihen. Um die Kundenanforderungen zufriedenstellend zu erfüllen und einen zuverlässigen Kundenservice zu bieten, musste Linde zusätzliche Lieferungen an die Kneipen durchführen, falls die Anforderungen der Kneipenbesitzer:innen in der ersten Lieferrunde nicht erfüllt wurden. Außerdem wurden die Zylinder in einigen Fällen unnötig herumgefahren und überflüssige Lieferungen an die Wirtshausbesitzer:innen vorgenommen. Bei der Analyse früherer Daten stellte Linde fest, dass 350.000 Gasflaschen pro Jahr unnötig herumgefahren und in verschiedene Teile des Vereinigten Königreichs geliefert wurden, wo Linde die Lösung testete.

KI-System:

KI-Algorithmen verwenden Auftragsverlaufsdaten und kombinieren diese mit Echtzeitdaten zu anderen externen Faktoren, um eine genauere Bedarfsprognose zu erstellen. Diese verbesserte Bedarfsprognose wird verwendet, um einen optimierten Lieferterminplan zu erstellen, der mit hoher Wahrscheinlichkeit an die Notwendigkeiten von Kund:innen angepasst wird. Ein solcher optimierter Zeitplan kann dazu beitragen, Überbestände und Unterbestände sowohl für Einkäufer:innen als auch für die Lieferfirmen zu reduzieren. Darüber hinaus können Algorithmen auf eine/n bestimmten Käufer/in zugeschnitten werden, indem die Bestellhistorie dieser Person genutzt und mit Echtzeitdaten zu anderen Faktoren wie lokalen Ereignissen, regionalen Feiertagen und den entsprechenden Wetterbedingungen kombiniert wird.

Das Digitalisierungsteam von Linde nutzte historische Daten zu Bestellinformationen von über 25.000 Kund:innen und verwendete KI, um den Einfluss anderer externer Faktoren wie Wetter, lokale Veranstaltungen, Feiertage, Lage der Kneipen und Sportveranstaltungen und deren Einfluss auf den Bierkonsum in Kneipen zu bestimmen, was sich dann auf die benötigte CO2-Menge auswirkte. Das Digitalisierungsteam erwähnte auch, dass es für jeden Kunden/jede Kundin (in diesem Fall Kneipenbesitzer:innen) einen „maßgeschneiderten Lieferalgorithmus“ haben könnte, der dazu beitragen würde, die richtige Anzahl von Zylindern zu liefern, die ein/e Kneipenbesitzer/in benötigt, und sie auch zum richtigen Zeitpunkt zu liefern.

Was ist unklar?

Die Ergebnisse des KI-Systems beeinflussen bzw. bestimmen die Routen und Fahrzeiten der Mitarbeitenden, die für die Auslieferung der CO2-Zylinder verantwortlich sind. Der Algorithmus basiert auch auf individuellen Verhaltensweisen und Eigenschaften von natürlichen Personen, jedoch auf denen von Kneipenbesitzer:innen und deren Kundschaft und nicht von Mitarbeitenden. Daher ist unklar, ob das KI-System eine Hochrisikoanwendung im Sinne von Anhang 3 Abs. 4 lit b) darstellt oder nicht.

Strafverfolgung

Vorhersage von Straftaten “im Auftrag”

ID 106

Unternehmensbereich: Buchhaltung und Finanzen

Name: Intelligente Risikoprüfung

Kontext:

Die Wirtschaftsprüfung versucht sicherzustellen, dass die Geschäftsbücher von Unternehmen ordnungsgemäß geführt werden, wie dies gesetzlich vorgeschrieben ist. Prüfer:innen prüfen die ihnen vorliegenden Aussagen, holen Nachweise ein und bewerten die Aussagen in ihrem Prüfungsbericht.

Prüfungswissen ist zu einem großen Teil implizites Wissen, das einzelne Fachkräfte durch Erfahrung erworben haben. Bei der Formulierung der Risikostrategie sind die Erkenntnisse einer Prüferin/eines Prüfers aus früheren Fällen sehr wertvoll. Deloitte wollte das implizite Prüfungswissen von Einzelpersonen für das gesamte Prüfungsteam zugänglicher machen. Um den Wissens- und Erfahrungsaustausch zu ermöglichen, begannen sie mit der Entwicklung des KI-Tools Guided Risk Assessment Personal Assistant, kurz GRAPA.

KI-System:

GRAPA unterstützt Wirtschaftsprüfer:innen dabei, eine gewählte Strategie gegenüber allen anderen zuvor verwendeten Risikostrategien abzugrenzen. Es verwendet eine Deloitte-Datenbank mit 10.000 Fällen, und jeder Fall enthält durchschnittlich fünfzig Risiken. GRAPA ist keine eigenständige Anwendung; vielmehr wird es der Software hinzugefügt, die Wirtschaftsprüfer:innen bei der Bestimmung der Risikostrategie verwenden. „Es ist, als ob Sie eine zweite Person bitten könnten, neben Ihnen zu lesen“, erklärt Van Gool (Audit Innovation Leader, Deloitte). „Aber der Vorteil ist, dass diese zweite Person über die gebündelte Expertise von Deloitte verfügt.“ Er betont, dass der/die Abschlussprüfer/in für die gewählte Risikostrategie und Prüfungsmethode verantwortlich bleibt. „GRAPA zeigt auf, was in ähnlichen Fällen passiert ist. Aber wenn die Situation eines Unternehmens etwas Besonderes oder Ungewöhnliches darstellt, liegt es natürlich am Prüfer/an der Prüferin, den Ansatz entsprechend anzupassen.

Was ist unklar?

Das KI-System wird im Kontext einer Wirtschaftsprüfung eingesetzt, um mögliche Risiken anhand ähnlicher Fälle zu erkennen. Unternehmen sind verpflichtet, eine ordnungsgemäße Steuererklärung vorzulegen und Verstöße können zu einer Straftat führen. Das KI-System hilft, derartige Verstöße zu vermeiden. Es erzeugt also eine Vorhersage, ob ein bestimmter Tatbestand zu einer Straftat führen kann.

Unklar ist, ob die Erkennung einer möglichen Straftat hier als „im Auftrag“ („on behalf“) einer durchsetzenden Behörde verstanden werden soll (vgl. Anhang 3 Abs. 6 lit a)), weil Unternehmen rechtlich verpflichtet sind, eine Steuererklärung zu erstellen.

ID 108

Unternehmensbereich: Buchhaltung und Finanzen

Name: Betrugsaufdeckung Beispiel 2

Kontext:

Die KI-basierte Überwachung von Transaktionen in Echtzeit kann Finanzinstitute bei der Bekämpfung der Geldwäsche und Zahlungsanbieter:innen bei der Aufdeckung von Betrug unterstützen. Durch Echtzeitzahlungen generierte Daten werden in das KI-System eingespeist, das dann verdächtige Transaktionen identifiziert, deren Verarbeitung stoppt und die Transaktion zur weiteren Überprüfung durch menschliche Compliance-Beauftragte markiert. Das Betrugserkennungssystem basiert auf KI-Algorithmen, die Muster erkennen und Verbindungen innerhalb der Daten identifizieren, die dann geclustert und klassifiziert werden. Mit der Zeit gewöhnt sich das System an die Daten und die Erkennungsgenauigkeit steigt.

KI-System:

Worldline, dem Erfolg und der Sicherheit seiner Kund:innen verpflichtet, führte A.S. Adventure zu einer innovativen Lösung – Fire by Fraugtser. Fire ermöglicht ein intuitives und dennoch ausgeklügeltes Schreiben von Regeln zur Betrugserkennung und übersetzt menschliche Denkprozesse in eindeutige Regeln. Darüber hinaus ermöglicht Fire Benutzer:innen, Regeln zu testen, bevor sie bereitgestellt werden, wodurch die Unsicherheit bei der Regelerstellung beseitigt und eine genaue Leistung sichergestellt wird. Der Wechsel zu Fire ermöglichte A.S. Adventure das einfache Schreiben und Testen von Betrugserkennungsregeln. Durch die Nutzung des KI-Scores von Fraugster konnte das Unternehmen Fehlalarme reduzieren, um gute und schlechte Kund:innen richtig zu identifizieren. Das Ausführen von Simulationen, bevor eine Regel live ging, ermöglichte es A.S. Adventure, zu erfahren, wie effektiv eine Regel sein kann. Die Regelleistung wurde verbessert, wodurch die Notwendigkeit manueller Überprüfungen entfällt und wertvolle Zeit beim Risikomanagement eingespart wird.

Was ist unklar?

Das KI-System unterstützt ein Geldinstitut bei der Umsetzung gesetzlicher Vorgaben zur Vermeidung von Geldwäsche. Unklar ist, ob die Erkennung einer möglichen Straftat hier als „im Auftrag“ („on behalf“) einer durchsetzenden Behörde verstanden werden soll (vgl. Anhang 3 Abs. 6 lit a)), weil Unternehmen rechtlich dazu verpflichtet sind, z. B. durch das Geldwäschegesetz.

ID 109

Unternehmensbereich: Buchhaltung und Finanzen

Name: Automatisierte Spesenprüfung

Kontext:

Die Automatisierungsherausforderung von Electrolux bestand darin, die Zentralisierung zu erhöhen und Prozesse zu verbessern und gleichzeitig den vielen Geschäftsreisenden ein nahtloses Erlebnis zu

bieten. Das Unternehmen musste auch seine hohen Standards und Ziele in Bezug auf die Einhaltung von Vorschriften und Richtlinien beibehalten. Electrolux prüfte 100 % der T&E-Ansprüche (Transport & Environment) manuell und stellte rechtzeitige, korrekte Erstattungen sicher – ein gründlicher, aber zeitaufwendiger und sich wiederholender Prozess. Spesenabrechnungen wurden zunächst von einem Manager genehmigt und dann im SSC Zeile für Zeile geprüft. Abgelehnte Anträge durchliefen den Prozess erneut, manchmal wiederholt. Quittungen erschienen in verschiedenen Sprachen und Berichte zeigten unterschiedliche Grade der Einhaltung der T&E-Richtlinien. Duplikate waren schwer zu erkennen, zusätzliche Genehmigungen verlangsamten den Betrieb und es war unmöglich, sich ein Gesamtbild von Wiederholungstäter:innen zu machen. Einige Prüfer:innen verfügten nicht über das Wissen und die Erfahrung, um alle Fehler und Anomalien zu finden, und es wurde zu viel Zeit für Ansprüche mit geringem Risiko aufgewendet, die gemäß den Richtlinien eingereicht wurden. Electrolux suchte nach einer Lösung, die den Prozess automatisieren und es seinen Prüfer:innen ermöglichen würde, sich nur auf T&E-Ansprüche zu konzentrieren, die ein höheres Maß an Aufmerksamkeit erfordern.

KI-System:

Electrolux hat lange nach einer innovativen Lösung gesucht, bevor es sich für AppZen entschieden hat. Expense Audit von AppZen konnte in das Spesenautomatisierungssystem von Electrolux integriert werden, um jeden Einzelposten in den Ausgaben in Echtzeit zu prüfen. Mit ihrer hohen Flexibilität kann das KI-System von AppZen Electrolux mit den wichtigsten Informationen aus Quittungen versorgen, um alle größeren Anomalien wie Duplikate, Ausgaben außerhalb der Richtlinien oder überhöhte Gebühren zu erkennen und die erforderlichen Richtlinienregeln einzuhalten. Das KI-System identifiziert selbstständig Einzelposten und deren Kostenarten und ordnet jede Transaktion dem zuständigen Mitarbeitenden zu. Dies verbessert die Durchsetzung von Compliance und Finanzvorschriften.

Was ist unklar?

Das KI-System unterstützt das Unternehmen bei der Einhaltung von Finanzvorschriften. Ein Verstoß dieser Vorschriften kann zu einer Straftat führen. Unklar ist, ob die Erkennung einer möglichen Straftat hier als „im Auftrag“ („on behalf“) einer durchsetzenden Behörde verstanden werden soll, weil Unternehmen dazu rechtlich verpflichtet sind (vgl. Anhang 3 Abs. 6 lit a)).

Bewertung von Dokumenten für Gerichtsverfahren

ID 54

Unternehmensbereich: Rechtswesen

Name: Intelligentes Vertragsmanagement

Kontext:

Um internationale Vorschriften einzuhalten, müssen Unternehmen mit Leasingverträgen Tausende von Verträgen einzeln durchgehen. Das ist eine immense Aufgabe, denn ein/e Analyst/in verbringt rund 90 Minuten mit jedem Vertrag. So mussten 2019 nach dem neuen Rechnungslegungsstandard IFRS 16 nahezu alle Leasingverträge bilanziert werden. Für ein Telekommunikationsunternehmen, das jeden Mast und jedes Grundstück, auf dem dieser Mast steht, pachtet, bedeutete dies, dass es Hunderttausende von Verträgen in allen möglichen Sprachen durchsehen musste.

KI-System:

Diese Zeit kann mithilfe von maschinellen Lerntechnologien drastisch reduziert werden. Um Unternehmen bei Aufgaben wie dieser zu unterstützen, hat eine Beratungsfirma eine benutzerfreundliche Anwendung entwickelt, die von Analyst:innen zur Überprüfung von Verträgen verwendet werden kann. Die Anwendung verfügt über einen Bot, der mit einer Reihe von Verträgen gefüttert werden kann. Der Bot gibt dem Analysten/der Analystin Vorschläge für Daten, die benötigt werden, z. B. das Startdatum eines Vertrags. Der/die Analyst/in sieht den hervorgehobenen Vorschlag und gibt an, ob er richtig ist oder nicht. Der Bot lernt daraus, was dazu führt, dass nachfolgende Verträge jedes Mal ein bisschen schlauer analysiert werden und die Zuverlässigkeit seiner Vorhersagen steigt.

Was ist unklar?

Das KI-System durchsucht Verträge nach bestimmten Inhalten und Fakten, die dann rechtlich bewertet werden, z. B. ob ein Vertrag verlängert werden muss, um eine finanzrechtliche Vorschrift einzuhalten. Unklar ist, ob im Fall eines Verfahrens, ein solches KI-System die Bewertung von Beweisen nach Anhang 3 Abs. 6 lit d) oder die Interpretation von Fakten nach Anhang 3 Abs. 8 lit a) übernimmt und somit eine Hochrisiko-Anwendung wäre. Die Ausnahme in Erwägungsgrund 40 (Entwurf von April 2021) für einfache administrative Aufgaben („ancillary administrative activities“) könnte Anwendung finden, was jedoch unklar ist, weil das KI-System für individuelle Fälle verwendet wird.

ID 109

Unternehmensbereich: Buchhaltung und Finanzen

Name: Automatisierte Spesenprüfung

Kontext:

Das Civil Rights Corps (CRC), eine gemeinnützige Organisation, die sich der Bekämpfung systembedingter Ungerechtigkeit innerhalb des amerikanischen Rechtssystems verschrieben hat, sah sich mit einem faktenreichen Fall mit mehr als 300.000 zu prüfenden Dokumenten konfrontiert. Angesichts mehrerer Angeklagter, einer komplizierten Reihe von Fakten und vielen Elementen, die bestätigt werden mussten, umfasste der Prozess der Faktensammlung die Überprüfung von Tausenden von Akten, um Hinweise aufzudecken, die beweisen würden, wie das private Bewährungssystem die verfassungsmäßigen Rechte ihrer Mandant:innen verletzt hat.

KI-System:

Die Ermittlungsfunktionen einer externen E-Discovery-Plattform ermöglichten es dem CRC-Team, Berge von Beweisen zu entdecken, indem es diese Dateien schnell durchsuchte. Sie nutzten eine spezielle Story-Building-Funktion, um die kritischsten Dokumente von drei Angeklagten zu verfolgen, virtuell zusammenzuarbeiten und sich erfolgreich auf Aussagen vorzubereiten. Durch die Optimierung des E-Discovery-Prozesses vom Hochladen und Verarbeiten von Daten bis hin zur Suche, Überprüfung und Produktion konnten sie aussagekräftige Informationen finden, verborgene Erkenntnisse ans Licht bringen und auf wichtige Beweise reagieren. Somit konnten sie ihrem Briefing-Antrag auf ein summarisches Urteil 95 Exponate beifügen. Seitdem hat das Team diese KI-Lösung in acht Fällen in sieben Bundesstaaten eingesetzt, um Akten zu durchsuchen, potenzielle Zeug:innen zu identifizieren und ihre einzigartigen Geschichten hervorzubringen.

Was ist unklar?

Das KI-System unterstützt das Finden, Verknüpfen und Zusammenfassen von Fakten für laufende

Prozesse und könnte somit unter Anhang 3 Abs. 6 lit d) oder Abs 8 lit a) fallen. In beiden Fällen ist unklar, ob die Nutzung „im Auftrag“ einer öffentlichen Stelle stattfindet.

Anhang II – Existierende EU Regularien

ID 101

Unternehmensbereich: Forschung und Entwicklung

Name: Produktentwicklung / Generatives Design

Kontext:

Die Produktion von Elektrofahrzeugen (EVs) bringt viele Herausforderungen mit sich. Obwohl die Automobilkonzerne ihnen gegenüber äußerst optimistisch sind – allein GM plant, bis 2023 mindestens 20 Elektro- oder Brennstoffzellenfahrzeuge auf den Markt zu bringen – sind solche Fahrzeuge teurer in der Herstellung. Für GM könnte generatives Design dabei helfen, diese Herausforderungen zu lösen, indem es leichtere Fahrzeuge und eine kürzere Lieferkette ermöglicht. Elektrifizierung und autonome Fahrzeuge werden die Automobilbranche grundlegend verändern. Daher ist es für Unternehmen künftig von entscheidender Bedeutung, in diesen hochtechnischen Bereichen eine Führungsposition einzunehmen.

KI-System:

In einer kürzlich durchgeführten Zusammenarbeit und unter Verwendung der generativen Designtechnologie entwarfen GM-Ingenieur:innen eine neue, funktional optimierte Sitzhalterung, ein Standardautoteil, das Sicherheitsgurtbefestigungen an Sitzen und Sitze an Böden sichert. Während die typische Sitzhalterung ein kastenförmiges Teil ist, das aus acht zusammengeschweißten Teilen besteht, hat die Software mehr als 150 alternative Designs entwickelt, die eher wie ein metallisches Objekt aus dem Weltraum aussehen. Das von GM gewählte Design besteht aus einem einzigen anstatt aus acht Edelstahlteilen, ist um 40 Prozent leichter und um 20 Prozent stärker als die vorherige Sitzhalterung.

Was ist unklar?

Elektrofahrzeuge sind Maschinen im Sinne der Maschinenrichtlinie (Directive 2006/42/EC), die in Anhang 2 Sektion A der KI-Verordnung zur Risikoklassifizierung gelistet ist. Die Gurthalterung könnte als sicherheitskritische Komponente des Fahrzeugs gelten, weil daran der Sicherheitsgurt angebracht ist. Jedoch wird das KI-System lediglich zur Entwicklung der Halterung verwendet, aber es ist nicht Teil des Fahrzeugs.

ID 89

Unternehmensbereich: IT und Sicherheit

Name: Analyse von Netzwerkbedrohungen

Kontext:

Ein globales Fortune-100-Pharmaunternehmen will den IT-Support für das Internet der Dinge (IoT), Wearables und andere nicht standardmäßige Geräte optimieren, die von seinen Geschäftsbereichen bereitgestellt werden. Um seine Branchenkund:innen besser bedienen und Gemeinkosten senken zu

können, muss das Unternehmen nicht standardmäßige mobile Geräte im gesamten Unternehmen sicher und effizient verwalten.

KI-System:

Das IT-Team verwendet ein Unified Endpoint Management (UEM) Tool eines externen Anbieters, um die Produktivität und Innovation im gesamten Unternehmen zu fördern und die Gemeinkosten zu minimieren. Endbenutzer:innen können den gesamten Self-Service-Anmeldungsprozess, von der Registrierung bis zum Herunterladen der App, innerhalb von 5 bis 15 Minuten abschließen. Das Tool hilft, das Unternehmen über potenzielle Endpunktbedrohungen auf dem Laufenden zu halten und Abhilfe zu schaffen, um Sicherheitsverletzungen und -unterbrechungen zu vermeiden. Die Lösung bot einen klaren Einblick in und Kontrolle über den globalen Gerätebestand des Unternehmens, der auf mehr als 80.000 unternehmens- und mitarbeitereigene Geräte und etwa 800 Apps angewachsen ist. Es half dem Team auch, Zeit zu sparen und Kosten zu senken, indem wichtige Konfigurations- und Supportprozesse automatisiert wurden.

Was ist unklar?

Es handelt sich hier um ein Pharmaunternehmen, sodass sich unter den betroffenen Endgeräten möglicherweise auch Medizinprodukte oder In-vitro-Diagnostikgeräte gemäß Anhang 2 Sektion A der KI-Verordnung befinden, z. B. Insulinpumpen. Unklar ist, ob ein KI-System mit dem Zweck, internetfähige Medizinprodukte oder In-vitro-Diagnostikgeräte vor Cyberangriffen zu schützen, als Sicherheitskomponente gilt.

ID 82

Unternehmensbereich: Logistik und Lieferketten

Name: Automatische Inspektion von Wirtschaftsgütern

Kontext:

In Industriezweigen wie der Logistik sind Schäden und Verschleiß an Betriebsmitteln im Laufe der Zeit an der Tagesordnung. Mithilfe einer Kamerabrücke zum Fotografieren von Güterzugwaggons sind KI-Systeme in der Lage, Schäden erfolgreich zu identifizieren, die Schadensart zu klassifizieren und die geeigneten Korrekturmaßnahmen zur Reparatur dieser Anlagen zu bestimmen.

KI-System:

Zunächst wurden entlang der Bahngleise Kameras installiert, um Bilder von vorbeifahrenden Waggons aufzunehmen. Die Bilder wurden dann automatisch in einen KI-basierten Bildspeicher hochgeladen, wo KI-Bildklassifizierer beschädigte Waggonkomponenten identifizierten. Die KI-Klassifikatoren wurden darin trainiert, wo sie in einem bestimmten Bild nach Waggonkomponenten suchen und wie sie erfolgreich Waggonteile erkennen und diese dann in sieben Schadensarten klassifizieren können. Als mehr Daten gesammelt und verarbeitet wurden, verbesserte sich die visuelle Erkennungsfähigkeit des KI-Systems in nur kurzer Zeit auf eine Genauigkeitsrate von über 90 %. Die vom System entdeckten Anomalien und Schäden wurden an ein Arbeitsplatz-Dashboard gesendet, das von Wartungsteams verwaltet wird.

Was ist unklar?

Das Bahnsystem fällt unter Anhang 2 Sektion B Abs (5 und das KI-System hat den Zweck, Schäden an

Waggons zu identifizieren und zu melden. Da unentdeckte Schäden sicherheitskritisch sein können, ist unklar, ob das KI-System eine Sicherheitskomponente im Sinne der KI-Verordnung darstellt.

ID 16

Unternehmensbereich: Produktion und Herstellung

Name: Prozesskontrolle und Optimierung

Kontext:

Linde ist ein globaler Anbieter von Industriegasen wie Stickstoff, Wasserstoff, Sauerstoff und vielen mehr und ist entlang der gesamten Wertschöpfungskette von der Produktion, Verarbeitung und Distribution bis hin zur Anwendung aktiv. Der Betrieb und die Steuerung von Gasverarbeitungsanlagen wirken sich sowohl auf die Produktivität als auch auf die Kosten aus, insbesondere auf die Energiekosten. Die Anlagensteuerung umfasst die Einstellung einzelner Kompressoren, Pumpen und Turbinen, Wärmetauscher und Ventile innerhalb einer Anlage, aber auch die Optimierung des Gesamtsystems einer Anlage als Ganzes.

KI-System:

Linde setzt künstliche Intelligenz ein, um das Verhalten der Anlage vorherzusagen und fein abgestimmte Strategien zur Reduzierung des Energieverbrauchs zu entwickeln. Das KI-System wird mittels Deep Learning in Kombination mit Reinforcement Learning implementiert. Das heißt, die Parameter der Anlage und ihrer Komponenten werden in einem neuronalen Netz abgebildet, das sich dann entsprechend einem vordefinierten Ziel des Algorithmus selbst optimiert. Dazu definieren Machine-Learning-Ingenieur:innen gemeinsam mit Fachexpert:innen dieses Ziel, die sogenannte Belohnungsfunktion (z. B. eine Reduzierung des Energieverbrauchs). Das KI-System wird in einer Anlage aufgebaut, die Sauerstoff und Stickstoff produziert und an eine/n direkt angeschlossene/n Kundin/Kunden liefert. Zuverlässigkeit und Reinheit müssen also jederzeit stabil sein. Linde konnte die Einstellungen einzelner Komponenten verfeinern, während die Anlage weiter mit stabiler Leistung lief.

Was ist unklar?

Das KI-System regelt die Energieversorgung der Gasverarbeitungsanlage, die vermutlich unter die Maschinenrichtlinie nach Anhang 2 Sektion A Abs 1 fällt. Möglicherweise enthält die Anlage auch Schutzsysteme zur bestimmungsgemäßen Verwendung in explosionsgefährdeten Bereichen (Anhang 2 Abs 6 - Richtlinie 2014/34/EU) oder Druckgeräten (Anhang 2 Abs 7 - Richtlinie 2014/68/EU). Unklar ist, ob die Energieversorgung einer solchen Anlage als sicherheitskritische Komponente anzusehen ist.

Diskussion: Ursachen für Unklarheiten

Auf Grundlage der KI-Systeme mit unklaren Klassifizierungen zeigt dieses Kapitel die darunter liegenden Ursachen auf, um entsprechende Verbesserungsvorschläge zu formulieren.

Kritische Infrastruktur

Definition von „Kritischer Infrastruktur“: Europäische oder nationale Definition?

Diskussion:

Zur korrekten Bestimmung von Hochrisiko-Systemen im Bereich der kritischen Infrastruktur ist essenziell, welche Definition von kritischer Infrastruktur zugrunde liegt: die nationale oder eine europäische? Mit dem Verständnis, dass Festlegung kritischer Infrastruktur eine nationale Kompetenz ist, wurde in dieser Analyse die nationale Definition des BSI herangezogen . Die Anwendung der nationalen Definition behindert jedoch das Ziel der KI-Verordnung, den Binnenmarkt zu stärken. Im Gegensatz könnte ein KI-System in manchen Mitgliedstaaten als hochrisikoreich klassifiziert werden, was zu einer Fragmentierung von Angebot und Einsatz beiträgt.

Der Entwurf des Europäischen Rats vom 25. November 2022 enthält eine Definition von kritischer Infrastruktur mit einem Verweis auf die Richtlinie über die Resilienz kritischer Einrichtungen 2022/2557, wonach die KI-Verordnung der europäischen Definition folgen würde (sofern der Vorschlag erfolgreich ist).

Vorschlag:

- Verwendung einer europäischen Definition von kritischer Infrastruktur zur Stärkung des Binnenmarktes.
- Stellungnahmen auf nationaler Ebene z. B. vom BSI in Deutschland, welche Definition bis zur Überführung der (neuen) Richtlinie 2022/2557 in nationales Recht anzuwenden ist.
- Erstellung einer Übersicht für KI-Anbieter, inwiefern sich die Definitionen von kritischer Infrastruktur der Mitgliedstaaten unterscheiden.

Anlagenarten und Schwellenwerte: Anpassungen für KI?

Diskussion:

Die nationale Definition des BSI spezifiziert Anlagenarten und Schwellenwerte, um KRITIS-Unternehmen zu identifizieren. Mit Blick auf die untersuchten KI-Systeme ist nicht klar, ob die Schwellenwerte entsprechend anwendbar sind. Zum Beispiel liegt bei einem „Intelligenten Verkehrssystem“ die „Anzahl angeschlossener Nutzer oder durchschnittlich im Versorgungsgebiet versorgter Nutzer“ bei 500.000 (vgl. Intelligente Verkehrssysteme Gesetz). Würde demnach ein KI-System zur Routenplanung, das auch sicherheitskritische Funktionen hat, als Hochrisiko-System gelten?

Vorschlag:

- Überprüfung und ggfs. Anpassung der Anlagenarten und Schwellenwerte für KI-Systeme durch nationale Stellen (z. B. BSI).
- Erhaltung und Stärkung der KI-Kompetenzen der nationalen Stellen, um KI-Anbieter bei der Anwendung der KI-Verordnung zu unterstützen, z. B. durch Beratung, Richtlinien oder im Rahmen von KI-Reallaboren (Regulatory Sandboxes).

Definition von „Sicherheitskomponente“ („safety component“)

Diskussion:

In einigen der untersuchten KI-Systeme ist unklar, ob das KI-System eine Sicherheitskomponente darstellt oder nicht. Diese Feststellung ist wichtig, weil das eine notwendige Bedingung für die Klassifizierung als Hochrisiko-System ist.

Der Entwurf des Europäischen Rats vom 25. November 2022 enthält in Erwägungsgrund 34 Beispiele von Sicherheitskomponenten im Bereich KRITIS („Examples of safety components of such critical infrastructure may include systems for monitoring water pressure or fire alarm controlling systems in cloud computing centres …“), die durchaus hilfreich sind.

Zur Erreichung der Klimaziele (vgl. European Green Deal) gibt es große Potenziale für den Einsatz von KI, wo jedoch häufig die Klassifizierung unklar ist. Zum Beispiel:

- Vorhersage des Verkehrsaufkommens in Städten und KI-gestützter Verkehrsfluss, etwa Ampelschaltungen
- Vorhersage des Energiebedarfs in Gebäuden und KI-gestützte Wärmeerzeugung in dezentralen Anlagen
- Vorhersage des Lebensmittelbedarfs und der Lebensmittelproduktion zur Verringerung von Abfällen

Vorschlag:

- Nennung weiterer Beispiele für KI-Sicherheitskomponenten in allen KRITIS Bereichen z. B. Wasser, Gesundheit, Ernährung.
- Klärung, ob eine Sicherheitskomponente auch reine Software sein kann.

Beschäftigung

Definition von „Aufgabe“ („Task“)

Diskussion:

In einigen Fällen erzeugt ein KI-System Aufforderungen für natürliche Personen, jedoch ist unklar, ob es sich dabei um eine Aufgabe im Sinne der KI-Verordnung handelt. Beispiele:

- Empfehlungen, etwa für eine Kochanleitung, z. B. wie man eine Ananas schält.
- Anweisungen bei der Navigation oder Routenplanung im Straßenverkehr.
- Aufträge, die eine mitarbeitende Person auszuführen hat, z. B. bei Lieferdiensten.

Der Entwurf des Europäischen Rats vom 25. November 2022 enthält in Erwägungsgrund 36 den Hinweis, dass KI-Systeme für den Zweck der Aufgabenverteilung Hochrisiko sein sollten, weil sie einen Einfluss auf die Karriereentwicklung haben.

Demnach sind mit „Aufgaben“ anscheinend Fälle gemeint, in denen eine angestellte Person eine Anweisung erhält, und die, bei Nichterfüllung, einen direkten, negativen und persönlichen Einfluss auf die berufliche Entwicklung dieser Person hat. Aufforderungen, die der/die Empfänger/in nach eigenem Ermessen befolgen kann oder für die es Alternativen gibt, oder bei Vorschlägen, die zur Inspiration dienen, wären dann keine Aufgabe im Sinne der KI-Verordnung. Insbesondere ist eine KI-generierte Aufforderung keine Aufgabe, wenn die Nichterfüllung keine negativen Konsequenzen für die betroffene Person hat.

Vorschlag:

- Ergänzung von Artikel 3 mit einer Definition für den Begriff „Aufgabe“.

Auslegung von „Arbeitsverhältnis“ („work-related contractual relationships“)

Diskussion:

Ein weiterer auslegungsbedürftiger Punkt im Kontext der Aufgabenverteilung durch KI ist die Beziehung zwischen Betreiber:in (User) der KI und der Person, die Aufgaben von der KI erhält (affected person), weil davon abhängt, ob bei Nichterfüllung negative Konsequenzen zu erwarten sind.

Varianten einer solchen Beziehung sind z. B.:

1. Ein Mitarbeiter eines Lebensmittelherstellers nutzt einen Koch-Chatbot für private Zwecke.
2. Ein Callcenter-Mitarbeiter bekommt Anfragen per KI zugewiesen.
3. Ein LKW-Fahrer bekommt Routenvorschläge durch eine KI.

Fall 1 fällt vermutlich nicht in die Hochrisiko-Klasse, weil bei Nichterfüllung keine negativen Konsequenzen zu befürchten sind. In Fall 2 und 3 ist die Klassifizierung als Hochrisiko-System denkbar, weil das Unternehmen mit dem Einsatz von KI eine Leistungssteigerung herbeiführen möchte. Entscheidend ist dabei, welche Erwartungen und Anweisungen der Arbeitgeber gegenüber dem Arbeitnehmer mit Blick auf das KI-System ausgesprochen hat. Falls der/die Beschäftigte eine

Alternative zum KI-System hat und dabei keine negativen Konsequenzen drohen, sind solche KI-Systeme möglicherweise Low-Risk.

Vorschlag:

- Klärung, in welchen Beziehungskonstellationen ein KI-System im Rahmen von Arbeitsverhältnissen als Hochrisiko-System einzustufen ist.
- Eine mögliche Konkretisierung der Hochrisiko-Klassifizierung ist z. B. wenn
  - der Arbeitgeber die Nutzung des KI-Systems autorisiert hat.
  - der Arbeitgeber den Arbeitnehmer angewiesen hat, den Empfehlungen bzw. Aufforderungen des KI-Systems zu folgen.
  - der Arbeitnehmer mit negativen persönlichen Konsequenzen rechnen muss, wenn er/sie dem KI-System nicht folgt.

Strafverfolgung

„Im Auftrag von“

Diskussion:

Die Kriterien in Anhang 3.6 gelten für Behörden und für Stellen, die im Auftrag von Behörden („on behalf“) handeln, jedoch ist unklar, unter welchen Umständen der zweite Tatbestand zutrifft. Wann erfolgt etwas „im Auftrag einer Behörde” und wann nicht?

Beispiele:

- Ein Finanzinstitut, das aufgrund gesetzlicher Anforderungen Maßnahmen zur Prävention von Geldwäsche durchführt
- Eine Organisation, die mit einem KI-System Zeugen vor Gericht hilft, „bessere“ Aussagen und Beweise vorzulegen
- Eine Anwältin, die eine Mandantin vor Gericht vertritt und KI-Systeme in der Vorbereitung verwendet

In diesen Fällen verlangt eine öffentliche Stelle (z. B. Finanzamt, Gericht) Informationen von betroffenen Unternehmen bzw. Personen (teilweise per Gesetz), aber im Einzelfall liegt nicht unbedingt eine direkte Beauftragung vor.

Der Entwurf des Europäischen Rats vom 25. November 2022 enthält in Erwägungsgrund 38 folgende Ausnahme:

„AI systems specifically intended to be used for administrative proceedings by tax and customs authorities as well as by financial intelligence units carrying out adminstrative tasks analysing information pursuant to Union anti-money laundering legislation should not be considered high-risk AI systems (...)“

Vorschlag:

- Klärende Ergänzungen in der KI-Verordnung, unter welchen Umständen der Einsatz von KI „im Auftrag einer Behörde“ stattfindet, z. B. in Erwägungsgrund 38
- Vollständige Auflistung bestehender Gesetze in der Finanzbranche, die ggfs. auch eine Ausnahme sind (neben dem Geldwäschegesetz).
- Aufbau von KI-Kompetenzen in den zuständigen Stellen auf nationaler Ebene (z. B. BaFin), um Unternehmen bei der korrekten Klassifizierung zu unterstützen.

„Mögliche Straftaten“ („criminal offence“)

Diskussion:

Anhang 3.6 a) beschreibt KI-Systeme zur Vorhersage von Straftaten, ob als Opfer oder Täter. Unklar ist, was genau als Straftat gilt bzw. gemeint ist, denn in den bisherigen Entwürfen der KI-Verordnung ist der Begriff (engl. criminal offence) nicht definiert. Laut dem deutschen Strafgesetzbuch (StGB) § 12 gilt jegliches rechtswidrige Verhalten als Straftat. Viele der untersuchten KI-Systeme (mit Blick auf den Einsatz in Unternehmen) unterstützen Aktivitäten, die gesetzlich geregelt sind, z. B.Straßenverkehr, Steuerrecht, Cybersecurity oder Arbeitsrecht. Hier muss klar abgegrenzt werden, unter welchen Bedingungen das KI-System die zentrale Rolle spielt, wann es eine unterstützende Funktion hat und wann der Mensch verantwortlich ist.

Ferner ist unklar, ob jeweils die nationale Definition von „Straftat“ Anwendung findet oder ob es eine einheitliche europäische Definition gibt.

Vorschlag:

- Definition von Straftat, „criminal offence“, in der KI-Verordnung, bzw. Verweis auf eine bestehende Definition auf EU-Ebene.
- Klärung unter welchen Bedingungen ein KI-System unterstützt und nicht als Hochrisiko-System einzustufen ist (vgl. adminstrative tasks)

Beweise und Fakten

Diskussion:

Anhang 3 6.d) und 8. a) beschreiben KI-Systeme zur Bewertung bzw. Interpretation von Beweisen und Fakten im Kontext eines Verfahrens. Diverse KI-Systeme dieser Studie unterstützen Aufgaben in diesem Bereich, z. B. das Durchsuchen von großen Textmengen nach relevanten Beweisen oder die Prüfung von Verträgen zur Erhebung bestimmter Fakten (z. B. Fristen).

Mit Blick auf die Risikoklassifizierung ist unklar, unter welchen Bedingungen ein Dokument (oder eine andere Information) als Beweis oder Fakt einzustufen ist.

Ist es zum Beispiel erforderlich, dass ein Verfahren bereits läuft oder sind auch Dokumente gemeint, die in der Zukunft ein Beweis werden könnten, etwa Verträge, die von einer KI-Anwendung überprüft

werden, um Fristen einzuhalten.

Vorschlag:

- Definitionen und Abgrenzung für die Begriffe „Beweis“ („evidence“) und „Fakt“ („fact“) in der KI-Verordnung.
- Erklärende Hinweise in den Erwägungsgründen 38 und 40

Anhang II

„Sicherheitskomponente“ in der KI-Verordnung

Diskussion:

Die Kriterien in Anhang 3.6 gelten für Behörden und für Stellen, die im Auftrag von Behörden („on Die Bestimmung, ob ein Produkt in den Anwendungsbereich der Regularien in Anhang 2 fällt, ist in der Regel problemlos möglich. Die Frage, ob ein KI-System eine Sicherheitskomponente (engl. „Safety Component“) ist, war in einigen der untersuchten Fälle unklar.

Zum Beispiel bei der Erkennung von Schäden an Zügen ist die KI weder Teil vom Zug noch von den Gleisen. Die Bilderkennung an sich (Kameras an Brücken) ist kein Teil eines Produktes oder des „Bahnsystems“ und stellt als solches auch kein Sicherheitsrisiko dar. Aber wenn das KI-System nicht oder fehlerhaft funktioniert, kann es zu Folgeschäden kommen.

In einem anderen Beispiel (nicht Teil der Studie) wird bei der vorausschauenden Wartung von Aufzügen eine KI-gestützte Smartphone-App verwendet, um durch die Sensoren im Smartphone vorherzusagen, ob ein Schaden am Aufzug besteht bzw. ob/wann eine Wartung nötig ist. Auch hier ist die App nicht Teil des Produktes (des Aufzugs), aber ein Ausfall oder eine Fehlfunktion der App kann zu Schäden führen.

Daher ist bei der Identifikation von Sicherheitskomponenten die Bestimmung der Systemgrenze von zentraler Bedeutung. Bei einer engen Auslegung (mit Fokus auf das Produkt nach Anhang 2), ist ein unterstützendes KI-System, z. B. zur vorausschauenden Wartung, möglicherweise keine Sicherheitskomponente.

Hinweis: Der Begriff „safety components“ ist im initialen Entwurf der Europäischen Kommission nicht definiert, aber sowohl das Parlament als auch der Europäischen Rat haben entsprechende Vorschläge mit der Definition veröffentlicht.

Vorschlag:

- Eindeutige Definition von „Sicherheitskomponente“ („safety components“) in der KI-Verordnung
- Klärung der anwendbaren Systemgrenzen bei der Identifikation von Sicherheitskomponenten (z. B. in den Erwägungsgründen).

„Sicherheitskomponente“ in den sektoralen Standards

Diskussion:

Der Begriff der „Sicherheitskomponente“ ist von zentraler Bedeutung bei der Risikoklassifizierung und neben der KI-Verordnung spielen auch die harmonisierten Standards eine große Rolle, weil sie den Begriff spezifizieren und in den sektoralen Kontext bringen.

Hier kann es zu Unklarheiten kommen, weil es bereits etablierte Standards gibt, die den Begriff „Sicherheitskomponente“ definieren, z. B. für Medizinprodukte im Kontext der essenziellen Funktionen (vgl. IEC 60601-1:2022) oder analog dazu in der Automobilindustrie (gemäß ISO 26262 und IEC 61508).

Zusätzlich gibt es nationale Gesetze und europäische Regularien, die ähnliche Definitionen enthalten (z. B. BSI-Gesetz §2 (13) für Kritische Komponenten in der Kritischen Infrastruktur, Anhang 3 in der RICHTLINIE 2014/33/EU über Aufzüge und Sicherheitsbauteile für Aufzüge).

Vorschlag:

- Bei der Bestimmung harmonisierter Standards durch CEN/CENELEC sollen die unterschiedlichen sektoralen Definitionen untersucht und bewertet und gegebenenfalls in Einklang gebracht werden.
- Kompetente Behörden in den Sektoren von Anhang 2 erhalten den Auftrag, Leitfäden für die Identifikation von KI-Systemen, die Sicherheitskomponenten in ihrem jeweiligen Sektor sind, zu veröffentlichen.
- Zertifizierungsstellen sind bezüglich der sektoralen Abweichungen bei der Definition von „Sicherheitskomponente“ sensibilisiert und erkennen diese im Rahmen der Konformitätsbewertung an.
- Die Varianten bei der Begriffsbestimmung (über sektorale Standards und nationalen Gesetzen) werden in der KI-Verordnung anerkannt, z. B. in Erwägungsgrund 34

Redundante Absicherungsmaßnahmen

Diskussion:

Nach der vorgeschlagenen Definition von „safety component“ im Entwurf der EU-Kommission (April 2021) können KI-Systeme eine Sicherheitskomponente sein, wenn deren Ausfall ein Risiko erzeugt:

> ‘safety component of a product or system’ means a component of a product or of a system which fulfils a safety function for that product or system or the failure or malfunctioning of which endangers the health and safety of persons or property;

In vielen Sicherheitskritischen Systemen gibt es bereits aufgrund bestehender Anforderungen redundante Sicherheitsmechanismen (z. B. der Kühlungskreislauf in Kernkraftwerken, die Stromversorgung in Krankenhäusern).

Daher ist unklar, ob eine KI-System mit einer Sicherheitsfunktion keine Sicherheitskomponente ist, wenn es redundante Sicherheitsmechanismen gibt. Beispiel: Bei der KI-gestützten Energieversorgung einer Gasverarbeitungsanlage gibt es möglicherweise redundante Systeme, die einspringen, wenn das KI-System ausfällt oder fehlerhaft ist. Wäre das KI-System in diesem Fall keine Sicherheitskomponente mehr?

Vorschlag:

- Anpassung der Definition von Sicherheitskomponenten in der KI-Verordnung mit Betonung auf Fälle, in denen das KI-System die primäre oder einzige Sicherheitskomponente ist.

Empfehlungen

Für die Politik

Bereich	Empfehlungen
Innovationen fördern	• Bereitstellung umfangreicher Leitfäden für die korrekte Risikoklassifizierung von KI-Systemen inklusive klarer Anweisungen und Beispielen, insbesondere bei KI in generischen und branchenunabhängigen Unternehmensbereichen. • Verbindliche und schnelle Beantwortung bei Fragen in Bezug auf unklare Klassifizierung über zentrales europäisches Portal (um unterschiedliche Interpretationen zu vermeiden) z. B. in Sandboxes. • Aufbau von Kompetenzen z. B. innerhalb der zuständigen Stellen auf Bundes- und Landesebene, um Firmen bei der der korrekten Risikoklassifizierung zu unterstützen. • Umsetzung von Kampagnen, die sich an alle Zielgruppen (Anbieter, Betreiber und Einzelpersonen) richten und über die Anforderungen und Obligationen der KI-Verordnung aufklären.
Kosten reduzieren	• Vereinheitlichung von Definitionen entlang von nationalen Gesetzen, europäischen Regularien und sektoralen Standards. Inkonsistente Definitionen erzeugen redundante Aufwände ohne einen Mehrwert zu schaffen. • Beschleunigung bei der Entwicklung von Standards und Leitfäden, die die Anforderungen der KI-Verordnung spezifizieren. Wartezeiten verzögern den Einsatz von KI durch rechtliche Ungewissheit. Mit Blick auf harmonisierte Standards empfiehlt sich der Zwang zu Third-Party-Zertifizierungen.

Anforderungen klären	Klare Klassifizierungsregeln schaffen Planungssicherheit bei Anbietern und Nutzern und verringern das Risiko einer fehlerhaften Klassifizierung. Definitionen und Auslegungen (Artikel 3 und Erwägungsgründen): • Welche Definition (inkl. Anlagenarten und Schwellenwerte) gilt für „Kritische Infrastruktur“, die europäische oder die nationale(n)? Die europäische Definition eignet sich besser zur Stärkung des Binnenmarktes und mit Blick auf die grenzüberschreitende Natur der Kritischen Infrastruktur. • Klärung der Definition von „Sicherheitskomponente“ (nur Hardware oder auch Software?) und die zu betrachtende Systemgrenze (z. B. mit Blick auf vorausschauende Wartung). • Klärung der Definition von „Aufgabe“ und der vertraglichen Beziehungskonstellation, bei der mit negativen Konsequenzen für den/die Arbeitnehmer/in gerechnet wird. • Klärung, wann ein KI-System „im Auftrag einer Behörde“ eingesetzt wird und welche Ausnahmen anwendbar sind (z. B. Geldwäschegesetz). • Klärung der Definition von „Straftat“ („criminal offence“) oder Verweis auf eine Definition auf EU-Ebene. Gleiches gilt für die Begriffe „Beweis“ („evidence“) und „Fakt“ („fact“).

Für Unternehmen

Bereich	Empfehlungen
Generelle Aspekte	• Die Konzentration auf KI-Systeme mit einem hohen strategischen Wert und einer niedrigen Risikoklasse kann sinnvoll sein, um die zusätzlichen Aufwände durch die KI-Verordnung gering zu halten. • Unternehmensbereiche mit tendenziell vielen Hochrisiko-Systemen sind das Personalwesen, der Kundenservice, Buchhaltung und Finanzen, IT und Sicherheit. • Unternehmensbereiche mit tendenziell wenigen Hochrisiko-Systemen sind Marketing und Vertrieb, Einkauf, Forschung und Entwicklung, Produktion und Herstellung, Logistik und Lieferketten. • Machen Sie sich mit der KI-Verordnung und den Klassifizierungsregeln vertraut (inkl. der Artikel, Erwägungsgründe und Anhänge). Nutzen Sie verfügbare Hilfsmittel von appliedAI oder ähnlichen Anlaufstellen.
Eigenentwicklung von KI-Systemen (Make AI)	• Führen Sie eine initiale Risikoklassifizierung zu einem frühen Zeitpunkt durch, um spätere Überraschungen zu vermeiden, weil sich durch die Klassifizierung als Hochrisiko-System die Kosten und die Komplexität erhöhen und ggfs. den strategischen Wert beeinflussen. • Beteiligen Sie bei der Risikoklassifizierung diverse Stakeholder, die idealerweise eine technische, juristische, kommerzielle und nutzerorientierte Sicht einbringen. • Lassen Sie, insbesondere bei größeren Investitionen, das Ergebnis der Risikoklassifizierung rechtlich bestätigen, um die Planungssicherheit zu erhöhen. Betrachten Sie dabei auch mögliche Varianten des KI-Systems in der Zukunft. • Berücksichtigen Sie die Risikoklasse und die anwendbaren Anforderungen über alle Phasen des KI-Lebenszyklus, insbesondere bei (vorhersehbaren) Veränderungen.

Nutzung von verfügbaren KI-Systemen **(Buy AI)**	• Fragen Sie den Anbieter des KI-Systems nach der zutreffenden Risikoklasse und ob diese auch in Ihrem Anwendungsszenario gleich bleibt. • Beachten Sie dabei, wer die Nutzung des KI-Systems freigibt, wer im Alltag damit arbeitet und wer von den Ergebnissen des KI-Systems betroffen ist. Wichtig ist, dass diese Personen frühzeitig informiert werden und sich ihrer Rolle bewusst sind. • Machen Sie sich mit den Anforderungen für Nutzer („User“ oder „Deployer“) der KI-Verordnung vertraut. • Besprechen Sie Änderungen des Einsatzszenarios vorab mit dem KI-Anbieter, da Änderungen eine Re-Klassifizierung bewirken können.

Empfehlungen

Wir haben diese Studie nach bestem Wissen und Gewissen durchgeführt und weisen daher ausdrücklich auf folgende Limitation hin:

Fokus auf KI im Unternehmen

Die untersuchten KI-Systeme sind alle dem Unternehmenskontext entnommen, d. h. KI in anderen Anwendungsbereichen, etwa für bestimmte Branchen (z. B. Medizin, Luft- und Raumfahrt, Automotive) oder Sektoren (z. B. Bildung, öffentliche Verwaltung, Gesundheitswesen), werden nicht berücksichtigt.

Aussagekraft der untersuchten KI-Systeme für Unternehmen in Europa

Die betrachteten KI-Systeme sind aktuell zwar in der Anwendung, jedoch nicht nur in Europa. Daher können wir keine Aussage darüber treffen, ob und inwiefern die Auswahl der KI-Systeme repräsentativ für KI in europäischen Unternehmen ist.

Begrenzte Informationen über KI-Systeme

Die Beschreibungen der KI-Systeme waren begrenzt und in manchen Fällen war der Mangel an Details ein Grund für eine unklare Klassifizierung. Mit mehr Informationen würde sich der Anteil unklarer Fälle möglicherweise verringern. Diese Beobachtung zeigt, dass für eine eindeutige Klassifizierung umfassende Details über das KI-System bekannt sein müssen.

Änderungen der KI-Verordnung

Diese Studie wurde während der laufenden Verhandlungen der KI-Verordnung angefertigt. Wir haben uns bemüht, für alle KI-Systeme den gleichen Maßstab anzulegen und abweichende Regeln aus neueren Entwürfen aufzugreifen und als solche anzuzeigen. Künftige Änderungen können zu anderen Klassifizierungen führen.

Mögliche Fehler bei der Ausarbeitung

Die KI-Verordnung ist ein umfassendes und komplexes Regelwerk und KI ist eine komplexe und facettenreiche Technologie. Beides entwickelt sich fortlaufend. Die Autoren haben sich mit beiden Themen umfassend beschäftigt und es gab diverse Review-Zyklen zur Überprüfung der Studie. Dennoch ist nicht auszuschließen, dass sich bei der Ausarbeitung Fehler eingeschlichen haben.

Autoren

Dr. Andreas Liebl

Dr. Andreas Liebl, Managing Director of appliedAI, is responsible for the appliedAI initiative. He serves as an expert regarding innovation and commercialization and as steering committee member in the Global Partnership on AI next to other advisory roles as in the Plattform Lernende Systeme or the KI Rat Bayern. In previous roles, he was a managing director of UnternehmerTUM, one of the largest innovation and entrepreneurship centers in Europe, he worked for McKinsey for five years and did his PhD at the Technical University of Munich.

Dr. Till Klein

Dr. Till Klein, Head of Trustworthy AI at appliedAI, is driving the acceleration of AI by co-creating methods, tools and insights in the field of AI Regulation and Governance to enable compliance and trust by adopters. He is a member of the OECD.AI and he has several years of industry experience in the regulatory roles, including Medical Devices, Drones, and as Quality Management System Auditor. Till is an Industrial Engineer and did his PhD on the evolution of collaboration networks in the context of technology transfer.

a.liebl@appliedai.de

t.klein@appliedai.de

Herausgeber

initiative for applied artificial intelligence

appliedAI Initiative Gmbh
www.appliedai.de

Freddie-Mercury-Straße 5
80797 Munich, Germany

Die appliedAI-Initiative ist Europas größte Initiative für die Anwendung modernster, vertrauenswürdiger KI-Technologie mit dem Ziel, die europäische Industrie zu Gestaltern im Zeitalter der KI zu machen und damit eine Welt zu schaffen, in der wir leben wollen. appliedAI agiert sowohl als Befähiger als auch als Innovator.

Bayerisches Staatsministerium für Digitales

Bayerisches Staatsministerium für Digitales
www.stmd.bayern.de

Oskar-von-Miller-Ring 35
80333 München

Danksagung

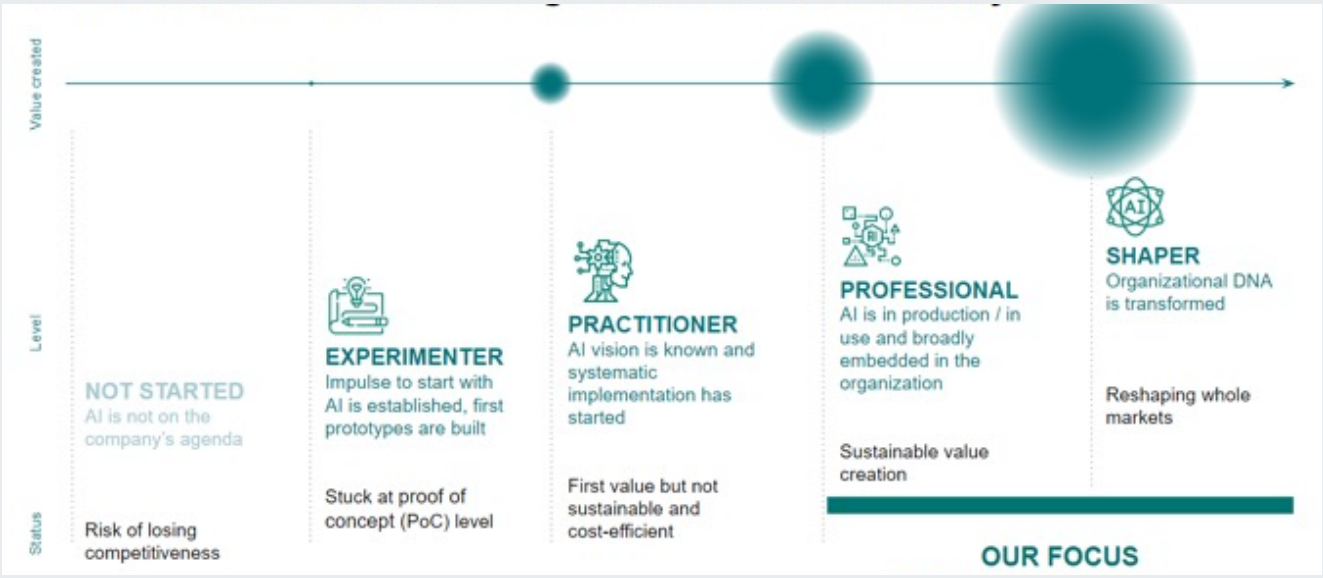
Die Autoren bedanken sich bei Dr. Tim Christiansen und Bianca Rabl vom Bayerischen Staatsministerium für Digitales für die gute Zusammenarbeit bei der Umsetzung und Publikation dieser Studie.

Ein herzlicher Dank geht auch an Rechtsanwalt Dr. David Bomhard und an Rechtsanwältin Jeannette Gorzala für ihr Feedback und ihre Anregungen.

Vielen Dank an das Team von appliedAI, das bei der Ausarbeitung der Studie geholfen hat, insbesondere Amela Gjishti, Susanne Klausing und Manuel Jimenez Medira.

Informationen zu appliedAI

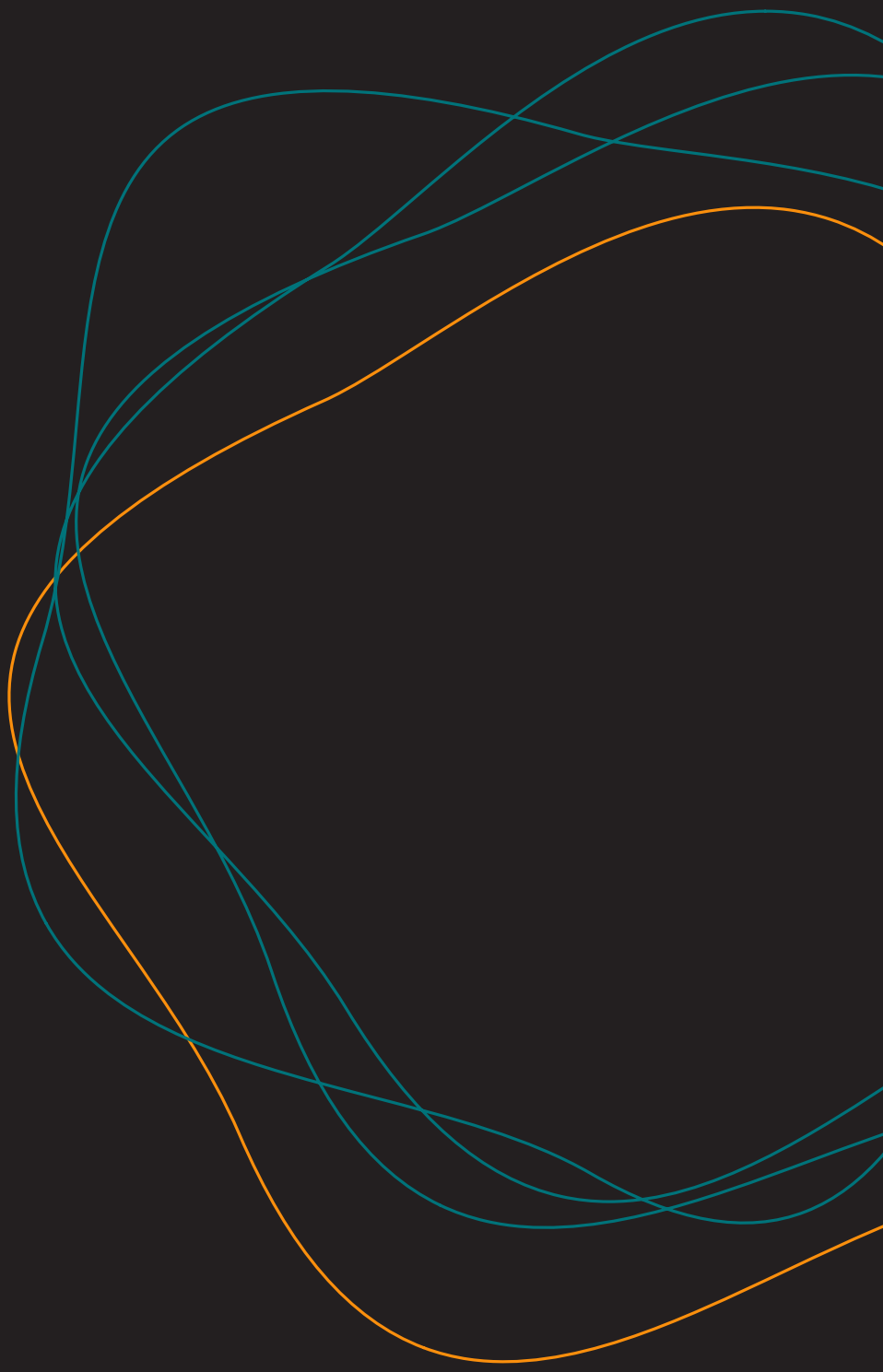
Die appliedAI Initiative wurde 2017 ins Leben gerufen und beschäftigt derzeit über 80 Mitarbeiter mit verschiedenen Hintergründen im Bereich der Künstlichen Intelligenz (KI). Ihr Ziel ist es, die europäische Industrie im Zeitalter der KI zu Gestaltern zu machen und damit eine Welt zu schaffen, in der wir leben möchten. Im globalen Rennen zur Technologieführerschaft kann dieses Ziel nur erreicht werden, indem wir zusammenarbeiten und voneinander lernen. Dabei konzentrieren wir uns darauf, Unternehmen in die professionelle KI-Anwendung zu begleiten, da nur dort wirklich Mehrwert entsteht.



appliedAI arbeitet mit Unternehmen, die eine Mentalität der Zusammenarbeit und Offenheit teilen, in Form von Partnerschaften daran, einzigartiges Wissen für die Anwendung vertrauenswürdiger KI zu schaffen, zugänglich zu machen sowie einen Austausch zu ermöglichen. Des Weiteren unterstützt die Initiative die KI-Transformation mit Lösungen und Dienstleistungen sowie umfassenden Programmen zur Beschleunigung der Einführung von KI.

appliedAI engagiert sich in diesem Kontext für die Wettbewerbsfähigkeit der europäischen Industrie unter Einhaltung zukünftiger regulatorischer Rahmenbedingungen und führt konkrete Aktivitäten, wie die Entwicklung eines KI-Risiko-Klassifizierungstools und den Aufbau einer MLOps-Infrastruktur, samt Werkzeugen und Prozessen, zur Compliance mit dem AI Act durch.

Unternehmen, die an einer Zusammenarbeit interessiert sind, können sich gerne melden.



White paper

appliedAI Initiative Gmbh
www.appliedai.de

August-Everding-Straße 25,
81671 Munich, Germany

initiative for
applied artificial
intelligence