



GENERATIVE KI MENSCHENZENTRIERT GESTALTEN UND NUTZEN

Informationen und Handlungsansätze
für Betriebs- und Personalräte

Impressum:

INPUT Consulting- Gemeinnützige Gesellschaft für Innovationstransfer,
Post und Telekommunikation mbH, Theodor-Heuss-Straße 2, 70174 Stuttgart
www.input-consulting.de

V.i.S.d.P.
Claus Zanker

Autoren:
Matthias Spielkamp
Lothar Schröder
Markus Hoppe
Karl-Heinz Brandl

Layout und Satz: Regine Lieb, klip GmbH
Zeichnungen: Tommi Süßmilch

ISBN:
978-3-910870-05-5

Stuttgart, Dezember 2025

**Diese und weitere Publikationen von INPUT Consulting finden Sie zum Download unter
<https://www.input-consulting.de/publikationen.html>**

GENERATIVE KI MENSCHENZENTRIERT GESTALTEN UND NUTZEN

**Informationen und Handlungsansätze
für Betriebs- und Personalräte**

INHALT

Generative KI als Neuland für Betriebspraxis und Mitbestimmung?	5
Was ist und kann generative KI?	10
Zwischen Entmachtung und Ermächtigung des Menschen: Welche Chancen, welche Risiken beinhaltet generative KI?	14
Welche Folgen hat generative KI für die Arbeit?	26
Wie lässt sich generative KI im Betrieb gestalten?	48
In welchen Rechtsrahmen ist der Einsatz generativer KI eingebettet?	56
Was ist beim Beschäftigtendatenschutz zu beachten?	70
Wer will noch generative KI mit welchen Mitteln begrenzen?	76
Autoren	82

GENERATIVE KI ALS NEULAND FÜR BETRIEBSPRAXIS UND MITBESTIMMUNG?

IT-Systeme mit generativer Künstlicher Intelligenz wie ChatGPT werden von Schüler*innen für Hausaufgaben genutzt, meistern Studienzulassungsverfahren und Prüfungen. Die Software generiert Texte, Programmcodes, Bilder und Videos und ist vielseitig einsetzbar. Generative Künstliche Intelligenz (KI) ermöglicht Anwendungen, die noch vor kurzer Zeit undenkbar schienen. Innerhalb von nur zwei Monaten nach Veröffentlichung im November 2022 hatte ChatGPT 100 Millionen aktive Benutzer, mehr als die Hälfte aller Deutschen haben den KI-Chatbot im November 2024 bereits genutzt. Im Juli 2024 berichtete die FAZ vom 2,9 Milliarden Besuchern von ChatGPT pro Monat.

Anwendungen mit generativer KI produzieren auf eine Texteingabe hin fotorealistische Bilder; sie benötigen nur eine wenige Sekunden lange Sprachaufzeichnung, um beliebige gesprochene Sprache in der Intonation des Sprechers zu erzeugen. Generative KI-Tools wie der SAP-Assistent Joule oder Microsoft Copilot versprechen, dass Arbeitsprozesse im Betrieb schneller, genauer und effizienter werden. Einer Umfrage des Digitalverbands Bitkom zufolge ist fast jedes zweite Unternehmen der Meinung, dass KI die Büroarbeit so revolutionieren wird, wie die Einführung des PC. In einer anderen Bitkom-Befragung zeigen sich knapp 80 Prozent der Unternehmen überzeugt, dass der Einsatz von KI künftig entscheidend für die Wettbewerbsfähigkeit der deutschen Industrie sein wird.¹

Ein Viertel der Beschäftigten nutzt generative KI bereits beruflich. Die Mehrheit davon hat positive Erfahrungen gemacht und damit Arbeitszeit gespart, hilfreiche Unterstützung bei schwierigen Aufgaben erhalten und wurden durch generative KI zu neuen Ideen inspiriert.² Gleichzeitig erzeugen sie damit oft Haftungsrisiken und nehmen nicht selten eine Verhaltenskontrolle durch Dritte genauso in Kauf wie Risiken für den Schutz der Unternehmensdaten.

1 <https://news.microsoft.com/de-de/generative-ki-in-der-industrie-hohe-erwartungen-aber-auch-noch-zoegern-bei-deutschen-unternehmen/>

2 Antonia Schlude u.a., Verbreitung generativer KI im privaten und beruflichen Alltag 2024, bidt, Dezember 2024

Betriebs- und Personalräte stehen angesichts der rasanten Verbreitung generativer Künstlicher Intelligenz vor einer großen Gestaltungsaufgabe. Dabei geht es darum, die Chancen derartiger Systeme für den betrieblichen Einsatz zu erschließen und gleichzeitig die Risiken einzudämmen. Dafür sind Kenntnisse von Technik, ethischen Werten, regulativen Stellschrauben, rechtlichen Grundlagen und Handlungsoptionen ebenso erforderlich wie die Notwendigkeit, zu verstehen, wo derartige Systeme an Grenzen stoßen oder nicht eingesetzt werden sollten. Dabei will diese Expertise helfen. Sie soll Betriebs- und Personalräte sowie weitere betriebliche Mitbestimmungsakteure dabei unterstützen:

- ☐ die Funktionsmechanismen, Einsatzmöglichkeiten, Wirkungen generativer KI zu verstehen,
- ☐ die Herausforderungen für die Mitbestimmungsorgane einzuordnen,
- ☐ Regulierungsoptionen und Rechtsgrundlagen kennenzulernen,
- ☐ Möglichkeiten der Einflussnahme zu erkennen,
- ☐ Gestaltungsanforderungen zu entwickeln,
- ☐ den Dialog unter den Mitbestimmungsakteuren zu fördern.

Die Gestaltung und Regulierung von Künstlicher Intelligenz hat zugunsten eines vertrauenswürdigen Einsatzes bereits begonnen. Die EU hat im Mai 2024 die KI-Verordnung verabschiedet, die nach einer Übergangszeit Rechtskraft in allen Mitgliedsländern entfalten wird. In der Presseerklärung dazu wird zu generativer KI ausgeführt: „Um dem breiten Aufgabenspektrum von KI-Systemen und der raschen Ausweitung ihrer Fähigkeiten Rechnung zu tragen, wurde vereinbart, dass KI-Systeme für allgemeine Zwecke (GPAI)³ und die GPAI-Modelle, auf denen sie beruhen, den ursprünglich vom Parlament vorgeschlagenen Transparenzanforderungen genügen müssen. Dazu gehört die Erstellung einer technischen Dokumentation, die Einhaltung des EU-Urheberrechts und die Verbreitung detaillierter Zusammenfassungen über die für das Training verwendeten Inhalte.“⁴

Der Hamburgische Beauftragte für Datenschutz und Informationsfreiheit empfiehlt: „Stellen Sie berufliche Chatbot-Accounts zur Verfügung. Beschäftigte sollten nicht eigenständig und unter Verwendung privater Daten ein Konto erstellen. [...] Formulieren Sie klare und dokumentierte interne Weisungen, ob, beziehungsweise unter welchen Voraussetzungen, welche Tools infrage kommen. Gegebenenfalls kann es auch sinnvoll sein, Betriebs- und Personalrät*innen mit ins Boot zu holen.“⁵

3 „General Purpose Artificial Intelligence“

4 Europäisches Parlament, Gesetz über künstliche Intelligenz: Einigung über umfassende Regeln für vertrauenswürdige KI, Pressemitteilung vom 09.12.2023

5 Der Hamburgische Beauftragte für Datenschutz und Informationsfreiheit, Checkliste zum Einsatz LLM-basierter Chatbots, 13.11.2023

Der Konzernbetriebsrat der Provinzial-Versicherung hat wahrscheinlich als erstes Mitbestimmungsorgan in Deutschland ein vorläufiges Regulativ zum Einsatz von ChatGPT vereinbart. Mit der Bezeichnung „Provinzial GPT Chat“ wurde im Feldversuch ein geschütztes und gekapseltes sogenanntes Sprachmodell ausprobiert. Gemeinsame Intention von Betriebsrat und Arbeitgeber war es, die Vorteile derartiger Systeme sowohl für das Unternehmen als auch für die Beschäftigten zu nutzen, die Risiken einzudämmen, aber auch Gestaltungserfahrungen zu sammeln.

Nach der getroffenen Vereinbarung sollte in den Modelltests u.a. Transparenz durch die Kennzeichnung maschinell generierter Inhalte geschaffen und kenntlich gemacht werden, wer diese Inhalte geprüft hat. Sogenannte Deep Fakes (realistisch wirkende Inhalte, die durch Systeme Künstlicher Intelligenz erzeugt, abgeändert oder verfälscht wurden) sollen als künstlich erzeugte Bilder gekennzeichnet werden. Die Nutzer*innen von „Provinzial GPT Chat“ sind in der Pflicht, den Output der Maschine auf Integrität, Genauigkeit und Zuverlässigkeit zu prüfen.

Über den Modelltest wurden Erfahrungen gewonnen, wie der Einsatz generativer KI-Systeme auf Qualifikationsbedarfe, Beschäftigung, Wertstatus und Arbeitserleichterung aus Sicht der Nutzer*innen und Betroffenen wirkt. Die Devise des KBR der Provinzial lautet: „Raus aus dem Staunen, rein in die Gestaltung“. Das ist eine Haltung, die die Autoren dieser Expertise unterstützen möchten.

Wir werden in dieser Broschüre nach einer Definition generativer KI (Kapitel 2) einen Überblick über deren Anwendungsmöglichkeiten geben und die Chancen und Risiken, die mit dem Aufkommen generativer KI verbunden sind, einordnen (Kapitel 3). Wie generative KI die Arbeit der Menschen verändert und welche Herausforderungen sich für die betrieblichen Interessenvertretungen daraus ergeben, ist Gegenstand des 4. Kapitels. Hier wird auch das Ergebnis einer Befragung von Betriebs- und Personalrät*innen zur Nutzung von generativer KI dargestellt. Auf Grundlage dieser Informationen geben wir in Kapitel 5 Praxistipps für Betriebs- und Personalrät*innen zur Gestaltung generativer KI im Unternehmen, bevor wir in Kapitel 6 (Alles, was Recht ist) einen Abriss geltender Gesetze und in Kapitel 7 (Wer noch mit welchen Mitteln generative KI begrenzen will) eine Übersicht über weitere Bestrebungen geben, generative KI zu regulieren.

WAS IST UND KANN GENERATIVE KI?



A cartoon illustration of two parrots perched on a horizontal branch against a solid red background. The parrot on the left is blue with a large, spiky crest and a white patch around its eye. The parrot on the right is red with a smaller crest and a white patch around its eye. Both have large, curved beaks. Above the blue parrot is a dark blue speech bubble, and above the red parrot is a dark red speech bubble. Both bubbles contain the text 'LISA MAG PFERDE & ÄPFEL' in white, bold, sans-serif capital letters. The blue parrot's tail feathers are long and pointed, while the red parrot's tail is shorter and more rounded.

LISA MAG
PFERDE & ÄPFEL

LISA MAG
PFERDEÄPFEL



WAS IST UND KANN GENERATIVE KI?

Seitdem ChatGPT im November 2022 für die allgemeine Öffentlichkeit zur Verfügung gestellt wurde, ist die so genannte generative KI in aller Munde und ein neuer Hype um KI ist entbrannt. Über die Gründe wird viel spekuliert; eine Ursache könnte sein, dass die Ausgabe von ChatGPT – die Generierung von Texten – für die meisten Nutzer*innen zugänglich, verständlich und nützlich ist. Denn auch vor der Veröffentlichung von ChatGPT konnten Menschen bereits bequem andere Systeme generativer KI nutzen, etwa Midjourney zum Erstellen von Bildern. Das hatte in erster Linie in Fachkreisen für Aufmerksamkeit gesorgt, etwa unter Fotograf*innen und Grafiker*innen.

In diesem Kapitel wird erklärt, wie generative KI funktioniert und welche Arten es gibt. Die Besonderheiten generativer KI im Vergleich zu anderen KI-Anwendungen werden erläutert und dargestellt, wie die Systeme bereits eingesetzt werden und werden könnten.

Definition

Als generative KI bezeichnet man IT-Systeme, die Inhalte – wie Texte, Bilder, Videos, Musik, Software – erstellen können, indem sie Muster in Trainingsdaten erkennen und Ausgaben erzeugen, die Daten nach ähnlichen Mustern neu zusammensetzen. Die Ergebnisse können menschengemachten Inhalten stark ähneln, obwohl sie nach einem statistischen Verfahren erzeugt werden. Im Gegensatz zu anderen KI-Anwendungen, die in erster Linie darauf ausgelegt sind, Daten zu analysieren oder zu interpretieren, zielt generative KI darauf ab, neue Daten oder Inhalte zu erzeugen.

Die populärsten Systeme generativer KI sind die sogenannten GPTs (**G**enerative **P**retrained **T**ransformers – auf Deutsch etwa: herstellende, vortrainierte Umwandler) wie ChatGPT. Aber nicht jede generative KI ist ein GPT. Systeme zum Erstellen von Bildern etwa beruhen auf GANs (Generative Adversarial Network- auf Deutsch etwa: herstellende kontradiktorische Netze). Daher tragen Stable Diffusion, Midjourney, DALL-E und andere Bilderzeuger GPT nicht im Namen.

Das vorausgeschickt, muss man sagen, dass es keine rechtsverbindliche Definition von generativer KI gibt. Die KI-Verordnung spricht nicht von generativer KI, sondern von „Mehrzweck-KI-Systemen“, die wiederum zirkulär definiert sind, als „KI-System, das auf einem Mehrzweck-KI-Modell basiert und das für eine Vielzahl von Zwecken eingesetzt werden kann, sowohl für die direkte Nutzung als auch für die Integration in andere KI-Systeme“.

Aber auch KI insgesamt ist ein Containerbegriff: Er umfasst ein sehr breites Spektrum an Bedeutungen, so dass es mitunter schwer fällt zu erkennen, wovon die Rede ist, wenn er verwendet wird. Daher hier eine kurze Liste weiterer wichtiger Begriffe:

☐ **Foundation Model** (Basismodell)

Bezeichnung für Modelle, die auf einer sehr großen Bandbreite an Daten trainiert wurden und dadurch für sehr viele Zwecke eingesetzt werden können. Der Begriff wurde von Forscher*innen aus Stanford geprägt und ist sehr kontrovers, da – mit guten Gründen – bestritten wird, dass es Eigenschaften gibt, die alle Modelle gemeinsam haben, die von der Definition umfasst sein sollen; zudem wird den Modellen wegen ihrer Unzulänglichkeiten abgesprochen, als generelle Grundlage für andere KI-Systeme und Anwendungen dienen zu können.

☐ **LLM** (Large Language Model)

Ein Modell, in den meisten Fällen ein Transformer (s. GPT), das mit sehr großen Textmengen trainiert wurde und dadurch extrem viele Beziehungen herstellen kann. LLM berechnen die statistische Wahrscheinlichkeit dafür, dass ein bestimmtes Wort auf ein anderes folgt. Die Kombination aus hochentwickelten statistischen Modellen und riesigen Mengen an Trainingsdaten führt dazu, dass LLM Texte erzeugen können, die Menschen so vorkommen, als seien sie von Menschen gemacht. Bekannte Modelle sind GPT-1 bis GPT-4 (OpenAI), LaMDA (Google), LLaMA (Meta), BLOOM (diverse Entwickler*innen).⁶

☐ **Multimodale Modelle, auch Multimodelle**

Systeme, die mehrere spezialisierte Modelle vereinen, um je nach Aufgabenstellung Ergebnisse zu liefern, etwa Bilder zu erkennen und passende Bildunterschriften zu verfassen. GPT-4 von OpenAI verfügt über multimodale Fähigkeiten.

Funktionsweise

GPTs errechnen Wahrscheinlichkeiten für Wortfolgen und erzeugen damit Texte. Wie sie das tun, kann hier nicht angemessen dargestellt werden; die Funktionsweise in vereinfachter Form allerdings schon. Wenn ein unvollständiger Satz lautet „Annalena Baerbock flog nach New“, kann das System die Wahrscheinlichkeit errechnen, mit der ein bestimmtes Wort diesen Worten folgt, und würde hier sehr wahrscheinlich „York“ ergänzen. Da es kein so genanntes Weltwissen hat und keinen Kontext kennt, könnte es damit falsch liegen, da Annalena Baerbock auch nach New Hampshire oder New South Wales hätte fliegen können. Diese Funktionsweise kann daher auch dazu führen, dass Systeme Aufgaben nicht lösen können, die Menschen trivial erscheinen, etwa wenn als Antwort auf „Was ergibt $2+2*5$ “ 20 statt 12 ausgegeben wird (weil das System keine Rechenregeln wie die „Punkt- vor Strichrechnung“ befolgt), oder Fehler passieren, die Menschen absurd erscheinen, etwa die Ausgabe „Angela Merkel ist deutsche Bundeskanzlerin“ (weil dieser Zusammenhang so häufig in den Trainingsdaten vorkommt).

In einem berühmt gewordenen Aufsatz haben die Forscherinnen Emily M. Bender, Timnit Gebru, Angelina McMillan-Major und Margaret Mitchell LLM daher „stochastische Papageien“ genannt, da sie lediglich Sätze ihren Wahrscheinlichkeiten entsprechend nachplappern.⁷

⁶ Liste von LLM, https://en.wikipedia.org/wiki/Large_language_model#List

⁷ On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? FAccT '21: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, March 2021 Pages 610–623, <https://dl.acm.org/doi/10.1145/3442188.3445922>

ZWISCHEN ENTMACHTUNG
UND ERMÄCHTIGUNG
DES MENSCHEN:

WELCHE CHANCEN,
WELCHE RISIKEN
BEINHALTET
GENERATIVE KI?





ZWISCHEN ENTMACHTUNG UND ERMÄCHTIGUNG DES MENSCHEN: WELCHE CHANCEN, WELCHE RISIKEN BEINHALTET GENERATIVE KI?

Anwendungsmöglichkeiten und Chancen

CHATGPT: Ein Programm stellt sich vor

„Vor ein paar Tagen bestand ChatGPT die Zulassungsprüfung für amerikanische Ärzte, dann die für Anwälte und schließlich auch noch die Diplomprüfung für den Master of Business Administration an der Ivy-League-Wirtschaftsuniversität Wharton School.“ Damit brachte uns ein Bericht in der Süddeutschen Zeitung im Januar 2023 ins Staunen.⁸ Dem Betreiber des Programms zufolge hat GPT-4 zuletzt eine Simulation einer Juraprüfung mit einem Ergebnis bestanden, welches ungefähr im Bereich der stärksten zehn Prozent der Teilnehmer*innen lag. Nach 5 Tagen knackte das Programm die Marke von einer Millionen-Nutzer*innen.⁹ Und nach zehn Monaten verleitet die Verbreitungsgeschwindigkeit und Anwendungsvielfalt von generativer KI einen Protagonisten der Entwicklung zu wichtigen Vergleichen: „Die sogenannte ‘Kambrische Explosion’“ bezeichnet einen kurzen Zeitraum vor mehreren Hundert Millionen Jahren, in dem es zu einer plötzlichen Vermehrung von komplexen Lebensformen kam. [...] Vor einem Jahr begann das, was mittlerweile oft als ‚zweite Kambrische Explosion‘ bezeichnet wird. Diesmal handelt es sich jedoch um künstliche Existenzformen, angetrieben durch das rasche Wachstum und die weitverbreitete Anwendung generativer KI.“¹⁰

Wir lesen und erfahren von bisher nicht für möglich gehaltenen Optionen, einen sehr niederschweligen Zugang zu programmtechnischen Möglichkeiten zu erhalten und „neue“ Bilder, Texte, Programmzeilen und Handlungsempfehlungen mit einer Geschwindigkeit zu erzeugen, die menschliche Fähigkeiten in den Schatten stellt. Aber was kann ChatGPT genau? Auf alle Fälle kann das Programm sich und seine Leistungsmerkmale vorstellen.

Über die Microsoft Suchmaschine Bing wurde ChatGPT nach den 25 am meisten genutzten Anwendungsmöglichkeiten gefragt. Daraufhin erstellte das Programm nachfolgende Tabelle. Diese wurde, nach einer zweiten Abfrage nach den Leistungsmerkmalen, vom Autor nur um die Möglichkeiten vervollständigt, Programmierarbeiten zu übernehmen. Als Quellen, um die Leistungsbeschreibung zu überprüfen, listet ChatGPT reichlich Webseiten auf.¹¹ Die Optionen,

8 Andrian Kreye, Denken mit System, Süddeutsche Zeitung vom 27.01.2023

9 Statista, Quellen-Business Insider, Unternehmensangaben

10 Ludwig Ensthaler, Wer an der KI wirklich verdient, FAZ 09. 10 .2023

11 <https://www.futurezone.de/digital-life/article439101/gpt-4.html>

<https://t3n.de/news/gpt4-chatgpt-openai-1541454/>

<https://www.heise.de/news/GPT-4-In-einer-Welt-rasender-KI-Entwicklung-Fakten-auf-einen-Blick-7546721.html>

<https://www.aicatcher.de/was-chatgpt-4-0-mit-plugins-kann/>

die das Programm bietet, dürften vergleichbar mit den Möglichkeiten anderer Anbieter generativer KI sein. Diese werden in der Zukunft wohl auch an Vielfalt gewinnen. Insofern ist der Vergleich mit der ‚Kambrischen Explosion‘ nicht falsch.

Abbildung 1: Anwendungsfelder von ChatGPT

Anwendungsfeld	ChatGPT 4.0 kann...
Chatbots	natürliche und flüssige Konversationen mit Menschen führen, indem Smalltalk gemacht wird, Fragen beantwortet oder Ratschläge gegeben werden.
Content-Erstellung	kreative und informative Inhalte wie Artikel, Blogs, Essays, Gedichte, Geschichten, Lieder, Witze oder Tweets generieren.
Übersetzung	Texte oder Sprache von einer Sprache in eine andere übersetzen, ohne dass ein vorheriges Training für jede Sprachpaarung erforderlich ist.
Sprachsynthese	Texte in natürliche und ausdrucksstarke Sprache umwandeln, die von verschiedenen Sprechern, Akzenten oder Emotionen gesprochen wird.
Spracherkennung	gesprochene Sprache in Text umwandeln, indem Hintergrundgeräusche reduziert und verschiedene Dialekte oder Slangs erkannt werden.
Bildgenerierung	realistische oder abstrakte Bilder aus Textbeschreibungen oder Skizzen erzeugen, z.B. Porträts, Landschaften oder Kunstwerke.
Bildbearbeitung	bestehende Bilder verändern, indem sie z.B. vergrößert, verkleinert, geschärft, geglättet, gefärbt, stilisiert oder restauriert werden.
Bilderkennung	Objekte, Gesichter, Szenen oder Emotionen in Bildern erkennen und beschreiben, z.B. für Sicherheit, Gesundheit oder Unterhaltung.
Videogenerierung	realistische oder abstrakte Videos aus Textbeschreibungen oder Bildern erzeugen, z.B. Animationen, Simulationen oder Filme.
Videobearbeitung	bestehende Videos verändern, z.B. schneiden, zusammenfügen, stabilisieren, verbessern oder vertonen.
Videoerkennung	Objekte, Gesichter, Szenen oder Emotionen in Videos erkennen und beschreiben, z.B. für Sicherheit, Gesundheit oder Unterhaltung.
Audio-Generierung	realistische oder abstrakte Audios aus Textbeschreibungen oder Bildern erzeugen, z.B. Musik, Geräusche oder Effekte.
Audiobearbeitung	bestehende Audios verändern, z.B. filtern, verstärken, mischen oder transkribieren.
Audioerkennung	Geräusche, Stimmen, Musik oder Emotionen in Audios erkennen und beschreiben, z.B. für Sicherheit, Gesundheit oder Unterhaltung.
Textanalyse	Texte auswerten, z.B. die Bedeutung, den Ton, die Absicht, die Stimmung oder die Qualität bewerten.
Textzusammenfassung	lange Texte in kurze und prägnante Zusammenfassungen umwandeln, z.B. für Nachrichten, Studien oder Berichte.
Textkorrektur	Texte auf Fehler oder Verbesserungsmöglichkeiten überprüfen und korrigieren, z.B. Rechtschreibung, Grammatik, Stil oder Klarheit.



Anwendungsszenarien und Einsatzmöglichkeiten

Texte erstellen

Die im Vergleich zu bisherigen Systemen sprachlich sehr guten Formulierungen lassen es attraktiv erscheinen, auch anspruchsvollere Texte von generativer KI erstellen zu lassen. Hier ist zu bedenken, dass vor allem überzeugende Formulierungen erkenntnisleitend sind. Nach dem Prinzip, „man weiß nicht, was man nicht weiß“ (und kann demnach faktische oder auch analytische Fehler nicht so gut – oder sogar gar nicht – erkennen), sollte man vor allem bei Themen, zu denen man keine Fachkenntnis besitzt, entweder auf Texte von generativer KI verzichten, oder von Kolleg*innen mit entsprechenden Kenntnissen prüfen lassen.

Übersetzen

Generative KI übersetzt bereits jetzt Texte mitunter besser als bisherige Systeme. Demnächst werden wir wahrscheinlich nicht mehr auswählen können, auf welcher Technologie die Übersetzungswerkzeuge basieren, die wir zur Verfügung haben. Einen qualitativen Sprung wird generative KI wohl bei der Übersetzung gesprochener Sprache ermöglichen: Schon jetzt können ausgewählte Podcasts bei Spotify automatisch in andere Sprachen übersetzt werden.¹² Gerade in Unternehmen, in denen Menschen unterschiedlicher Herkunft arbeiten, kann die Übersetzung von Trainingsvideos oder auch die Live-Übersetzung von Gesprächen – sowohl per Telefon als auch in Videokonferenzen – einen sehr großen Unterschied bedeuten. Doch auch hier ist weiterhin Vorsicht geboten. Die Kombination daraus, dass generative KI wenig Trainingsdaten so genannter „kleineren“ Sprachen zur Verfügung hat, mit der Tatsache, dass die Systeme immer wieder Aussagen „erfinden“, stellt auch bei Übersetzungen ein Risiko dar.

Recherchieren

Viele Anbieter*innen, darunter Microsoft und Google, integrieren bereits generative KI in ihre Internet-Suche, etwa in Form von Microsoft Copilot. Zahlreiche Studien belegen, dass die Ergebnisse nicht nur oft falsch sind. Durch die Darstellung, in der auf Quellen verwiesen wird, wird zusätzlich der Eindruck von Zuverlässigkeit erweckt, die jedoch nicht gegeben ist. Oft genug werden die Informationen der Quellen, auf die verwiesen wird, falsch dargestellt. Somit kann diese Angabe von Quellen sogar das Risiko erhöhen, da falsche Informationen als glaubwürdig wahrgenommen werden. Generative KI kann allerdings, neben der Internetsuche, auch dafür verwendet werden, die Suche in unternehmenseigenen Datenbeständen zu verbessern. Auch hier ist die Gefahr von Falschdarstellungen nicht beseitigt.

Bilder erstellen

Generative KI kann dazu genutzt werden, Bilder in allen Stilarten zu erzeugen – von fotorealistischen Abbildungen bis hin zu Bildern, die wie (fotografierte) Ölgemälde aussehen. Sollen Systeme dafür verwendet werden, sind Fragen zu Rechten, De-Skilling, Auswirkungen auf andere Jobs, aber auch Haftungsfragen zu beachten – etwa, wenn auf Bildern geschützte Marken verwendet werden oder eine erkennbare Person in einem Zusammenhang dargestellt wird, in dem sie nicht erscheinen möchte.

Software erstellen

Viele Beobachter*innen sehen im Programmieren eine der wichtigsten Einsatzmöglichkeiten von generativer KI. Schon ist von so genannten „citizen developers“ die Rede, von Bürger-Softwareentwickler*innen, die mithilfe von generativer KI ausgebildete Programmierer*innen ersetzen. Einige Expert*innen bremsen derzeit die Erwartungen und gehen davon aus, dass viele komplexe und zentrale Programme weiterhin von erfahrenen Profis entwickelt werden.¹³ Diese können dann aber modularartig auch von weniger versierten Menschen mithilfe von generativer KI kombiniert und für Zwecke eingesetzt werden, die nicht unternehmens- oder sicherheitskritisch sind. Ob und wenn ja, wie schnell sich diese Einschätzung ändern wird, ist nicht abzusehen.

12 <https://newsroom.spotify.com/2023-09-25/ai-voice-translation-pilot-lex-fridman-dax-shepard-steven-bartlett/>

13 <https://www.heise.de/news/Drei-Fragen-und-Antworten-KI-ist-nichts-fuer-Citizen-Developer-9539316.html>



Ideenentwicklung

Oft genug ist ein weißes Blatt Papier die größte Hürde im Kreativprozess. Generative KI wird auf den Prompt „Gib mir zehn Ideen zum Thema x“ auf jeden Fall zehn Vorschläge machen – egal, wie gut oder schlecht. Das kann dabei helfen, diese Anfangshürde zu überwinden.

Datenanalyse

Generative KI kann Daten analysieren und komplexe Modelle und Simulationen erstellen, um neue Erkenntnisse zu gewinnen. Das kann den Forschungsprozess beschleunigen, muss aber mit derselben Vorsicht betrachtet werden wie andere Anwendungen. So wurde beispielsweise KI dazu genutzt, Vorschläge für neue Materialien zu entwickeln. Bei näherer Betrachtung erwiesen sich die Molekülstrukturen vieler dieser vermeintlich neuen Materialien als nicht umsetzbar.¹⁴

Beispielhafte Anwendungsszenarien: Wo generative KI im Betrieb ihren Zweck erfüllt

Banken

„Wall-Street-Banken gehen rigoros gegen die Verwendung von ChatGPT vor, berichtet Bloomberg. Eine Reihe globaler Investmentbanken wie Bank of America, Citigroup, Deutsche Bank AG, Goldman Sachs und Wells Fargo führten demnach Beschränkungen für die Technologie ein.“¹⁵

In Deutschland zeigt sich der Bankenverband zwar skeptisch beim Einsatz des Systems im direkten Kontakt mit Kund*innen. Verlockend scheinen die Anwendungspotenziale trotzdem. Der Verband zieht Anwendungsfelder in Betracht, „die ein größeres Textverständnis und kreative Formulierungsfähigkeiten erfordern, wie zum Beispiel die Erstellung von internen Analysen und Zusammenfassungen oder das Entwerfen von Inhalten und Texten für Produktbeschreibungen oder Werbekampagnen in den Bereichen Marketing und Vertrieb. Unterschiedlichste Arbeitsaufträge wie eine Recherche oder das Erstellen eines Strukturierungsvorschlages könnten zukünftig an die Sprach-KI gerichtet werden und das in natürlicher Sprache, quasi wie eine E-Mail an einen Mitarbeitenden. Die Sprachfähigkeiten generativer KI-Modelle könnten auch für Programmieraufgaben genutzt werden, womöglich sogar zur Pflege oder Migration von bestehenden Legacy-Systemen auf aussterbenden Programmiersprachen wie Cobol, die heute nur noch wenige IT-Fachkräfte beherrschen.“¹⁶

Versicherungen

Die spezifischen Merkmale des Geschäftsmodells „Versicherung“ wirken heute als Unterstützungsfaktoren für eine Verbreitung von KI-Systemen in der Branche. Es werden immaterielle Produkte vertrieben, große Datenmengen sind verfügbar, ein dezidiertes Risikomanagement ist bereits vorhanden und mathematisch fundierte Analyseerfahrung ist in den Firmen verankert. Die Geschäftsprozesse sind im Branchenvergleich stark regelbasiert und standardisiert. Das Management hat Erfahrung mit hoher Regulierungsintensität und die Branche nimmt eine Vorreiterrolle in der Digitalisierung ein.¹⁷

14 https://x.com/Robert_Palgrave/status/1730358675523424344

15 <https://www.golem.de/news/sicherheitsbedenken-banken-verbieten-verwendung-von-chatgpt-2302-172191.html#:~:text=-Sicherheitsbedenken%20Banken%20verbieten%20Verwendung%20von,und%20ähnliche%20Tools%20zu%20verwenden.&text=Wall%2DStreet%2DBanken%20gehen%20rigoros,von%20ChatGPT%20vor%2C%20berichtet%20Bloomberg.%; letzter Zugriff 18.12.23>

16 <https://bankenverband.de/digitalisierung/chatgpt-wie-nuetzlich-ist-die-generative-kuenstliche-intelligenz-fuer-banken/>; letzter Zugriff 18.12.23

17 Enquetekommission „Künstliche Intelligenz- gesellschaftliche Verantwortung und wirtschaftliche, soziale und ökologische Potenziale“ EKKI, S. 153

Der Gesamtverband der Versicherer nennt zahlreiche Anwendungen von KI-Systemen, die im deutschen Versicherungsmarkt bereits eingesetzt werden, und hebt die Vorteile aus Kundensicht hervor:

„Kunden oder die Autowerkstatt können bei Schadenfällen im Kfz-Bereich durch vollautomatisierte Prozesse von der Einreichung der Schadensbilder via App bis zur Bewilligung der Schadenssumme deutlich schneller ihr Geld erhalten.

Roboadvice, Chatbots und Sprachassistenten sind rund um die Uhr für Kunden verfügbar und reagieren mit schnelleren Reaktionszeiten auf Serviceanfragen, bei einer hohen einheitlichen Beratungsqualität.

Die Versichertengemeinschaft kann noch besser vor Betrügereien geschützt werden, etwa durch intelligente Bilderkennung, die Fälschungen z.B. bei vermeintlichen Kfz-Schäden besser und schneller erkennt.“¹⁸

Medienbranche

Schon vor dem Durchbruch von ChatGPT gab es zahlreiche Anwendungsszenarien für KI in der Medienbranche. 2021 beschrieb Andrian Kreye in der Süddeutschen Zeitung zusammenfassend einige Praxisbeispiele: „69 Prozent der befragten Entscheidungsträger sehen künstliche Intelligenz als wichtigsten Treiber für Innovation im Journalismus. Manche Firmen sind beim Einsatz von KI in der journalistischen Arbeit auch schon relativ weit. Bloomberg News gibt zum Beispiel an, rund ein Drittel seiner Nachrichten mithilfe von KI zu generieren und zu verbreiten. Auch die Süddeutsche Zeitung verwendet solche Technologien, wenn auch nicht in diesem Ausmaß. [...] Die BBC entwickelte einen Chatbot (also einen automatisierten Gesprächspartner), der die unzähligen Nutzerfragen zum Coronavirus auf der Basis seriöser Informationen automatisierte. Andere Medienfirmen nutzen KI, um das Ausspielen von Inhalten auf Interessen und Bedürfnisse zuzuschneiden, auf ihren Webseiten potenzielle Abonnenten zu identifizieren oder ihre Archive zu sortieren und zugänglicher zu machen.

Es gibt auch sogenannten Robojournalismus. Das sind Texte, die ohne menschliche Hilfe von KI-Anwendungen verfasst werden. Allerdings sind die Themenbereiche sehr begrenzt. Letztlich funktioniert das nur mit Geschichten, die auf Zahlen basieren. Wirtschaftsmeldungen zum Beispiel, Sport- und Wahlergebnisse.

Die Los Angeles Times wiederum hat eine KI namens Quakebot entwickelt, die aus seismografischen Daten Nachrichten über Erdbeben generieren kann. Berühmt wurde der Algorithmus, weil er 2014 einen Bericht über ein kalifornisches Beben schon drei Minuten nach dem Ende der Erschütterungen lieferte.

Des Weiteren gibt es Algorithmen, die den „Bias“ in Medienorganisationen ermitteln, also versteckte Tendenzen und Vorurteile. Das kann sowohl der unbewusste Rassismus in Beiträgen sein also auch die ungleiche Verteilung von Themen. Das Aijo Project der London School of Economics and Political Science hat so eine Methodik entwickelt, die schon acht große Nachrichtenorganisationen wie Reuters, Agence France Presse und die Deutsche Welle einsetzen.“¹⁹

18 GDV-Gesamtverband der Versicherer, Positionspapier des GdV „Verwendung von Algorithmen in der Versicherungswirtschaft vom 10.05 2019“

19 Andrian Kreye, SZ Künstliche Intelligenz: Die rote Linie, 15.02.2021



Am 8. August 2023 ging das KI- Webradio „bigGPT“ auf Sendung. Eine virtuelle Moderatorin präsentiert ein tagesaktuelles Musikprogramm und chattet mit den Zuhörer*innen über KI-Themen. „Die Trainingsdaten für die Moderatorin basieren auf den GPT-Modellen sowie Prompts von Redakteur*innen, die der Software tagesaktuelle journalistische Aufgaben für die aktuelle Berichterstattung und zur Musikzusammenstellung geben. [...] bigGPT ist komplett aus Computer generierten Inhalten und synthetischen Stimmen hergestellt, ein sogenanntes Large Language Modell.“²⁰

Grenzen und Risiken

Kaum eine andere Technologie wird derzeit so kontrovers diskutiert wie generative KI. Die Gründe dafür sind vielfältig; in einem Überblick werden hier die Wichtigsten vorgestellt. Wir beschränken uns auf Szenarien, in denen Systeme zu legitimen Zwecken genutzt werden, in denen also Probleme für die Anwender*innen entstehen – nicht solche, in denen sie verwendet werden, um Schaden anzurichten, etwa indem mit ihrer Hilfe Desinformation oder Schadsoftware erzeugt werden.

Generative KI-Systeme produzieren häufig falsche Aussagen und Inhalte, die sehr überzeugend aussehen

Da die Systeme weder eine Vorstellung von Wahrheit oder Wahrhaftigkeit haben noch Absichten verfolgen, wäre es falsch zu schreiben, dass sie lügen. Aber ihre Ausgaben sind häufig falsch. Was genau „häufig“ bedeutet, hängt vom Kontext ab. So kam eine Untersuchung von Bing Chat, dem auf ChatGPT basierenden Chatbot der Microsoft-Suchmaschine Bing, zu dem Ergebnis, dass ein Drittel der Antworten von Bing Chat auf Fragen zu den Wahlen in Bayern, Hessen und der Schweiz im Herbst 2023 sachliche Fehler enthielt. Zu den Fehlern gehörten falsche Wahldaten, veraltete Kandidat*innen oder sogar erfundene Skandale über Kandidaten*innen.²¹ Das ist bereits für sich genommen ein großes Problem, wiegt aber gerade deshalb besonders schwer, weil die Aussagen sehr überzeugend formuliert sind. Zwischen wahren und falschen Ausgaben ist sprachlich kein Unterschied zu erkennen.

Eine Studie von Informatikern der Université du Québec in Kanada hat ergeben, dass ChatGPT 3.5 dazu neigt, unsicheren Code zu generieren. Die Forscher kommen zu dem Schluss, dass die KI, obwohl sie eigentlich Kenntnisse über Programmierschwachstellen besitzt, Code generiert, der nicht robust und dadurch unsicher ist.²²

20 <https://www.radioszene.de/178328/bigpt-bigfm-ki-radio.html>; letzter Zugriff 18.12.2023

21 Riccardo Angius/Raziye Buse Çetin/Marc Faddoul/Alex Felder/Clara Helming/Waldemar Kesler/Natalie Kerby/ Alexandra Kimel/ Michele Loi/Angela Müller/Melis Omalar/Estelle Pannatier/Simone Robutti/Salvatore Romano/Isabel Sanchez/Anna Lena Schiller/Miazia Schueler/Denis Sorie/Matthias Spielkamp/Marc Thümmeler/Tobias Urech/Mira Zimmermann, Generative AI and elections: Are chatbots a reliable source of information for voters? An analysis of Microsoft's Bing Chat, December 2023, https://algorithmwatch.org/de/wp-content/uploads/2023/12/AlgorithmWatch_AIForensics_Bing_Chat_Report.pdf

22 Raphaël Khoury/Anderson R. Avila/Jacob Brunelle/Baba Mamadou Camara, How secure is Code generated by ChatGPT? Quebec 2023, <https://arxiv.org/pdf/2304.09655>

Zur Herstellung und Verwendung von generativer KI werden enorme Ressourcen verbraucht

Hier geht es vor allem um Strom und Wasser. Wie groß der Verbrauch ist, kann derzeit nur geschätzt werden, da Entwickler*innen und Anbieter*innen nicht dazu verpflichtet sind, entsprechende Daten offenzulegen. Anbieter der Systeme geben gerne nur an, wie hoch der direkte Energieverbrauch und die Emissionen für einen Trainingszyklus sind. Durch den Energieverbrauch während der Trainingsphase wurden etwa beim BLOOM-Modell 24,7 Tonnen CO₂-Äquivalente emittiert. Wenn aber die Hardware-Produktion und die Betriebsenergie in die Rechnung einfließen, verdoppelt sich der Emissionswert bereits. Darin sind allerdings noch nicht die kontinuierlichen Emissionen während der Anwendung des Modells enthalten. Es fehlen verlässliche Zahlen aus dieser sogenannten Inferenzphase, aber erste Indizien weisen darauf hin, dass die Emissionswerte sehr hoch sind – sowohl bei der Herstellung der notwendigen Hardware für die Anwendung als auch bei ihrem Betrieb.²³

ChatGPT benötigt 500 Milliliter Wasser für einen einfachen Dialog mit 20 bis 50 Fragen und Antworten. Da der Chatbot mehr als 100 Millionen aktive Nutzer*innen hat, von denen alle mehrere Dialoge eröffnen, ist der Wasserverbrauch von ChatGPT schwindelerregend. Und diese Zahlen beziehen sich nur auf den Anwendungsmodus: Das Training von GPT-3 in den State-of-the-Art-Rechenzentren von Microsoft in den USA kann direkt 700.000 Liter sauberes Frischwasser verbrauchen, was für die Produktion von 370 BMWs oder 320 Teslas ausreichen würde. Der Wasserverbrauch beim GPT-3-Training würde sich sogar verdreifachen, wenn es in den asiatischen Microsoft-Rechenzentren stattfände. Bisher werden solche Informationen aber geheim gehalten.²⁴ Ohne eine umfassende Lebenszyklusanalyse können wir den ökologischen Fußabdruck von KI-Modellen nicht angemessen erfassen, müssen aber davon ausgehen, dass er sehr groß ist.

Für das Training von Modellen arbeiten Menschen unter unwürdigen Bedingungen

Diese Tatsache ist nicht auf generative Systeme beschränkt; schon lange vorher haben Menschen bei sehr schlechter Bezahlung und unter großen psychischen Belastungen Daten für das Training von KI-Modellen sortiert und etikettiert („gelabelt“). Doch dabei ging es meist darum, dass die Arbeit eintönig war und unter großem Zeitdruck erledigt werden musste. Bei generativer KI kommt hinzu, dass die Systeme „toxische“ Inhalte ausgeben, beispielsweise lebhaftes Darstellungen von sexueller Gewalt, Suiziden, Tierquälerei und mehr. Diese Beschreibungen müssen von Menschen gelesen und eingeordnet werden, damit sie, bevor das Modell vielen Nutzer*innen zur Verfügung gestellt wird, aussortiert werden können. Das machten im Fall von OpenAI Menschen in Kenia zu einem Stundenlohn zwischen 1,25 und 2 US-Dollar. Später sollen dann auch Bilder überprüft worden sein, die zum Teil Gewalt an Kindern, Morde und Vergewaltigungen darstellten.²⁵

23 Anne Mollen/Kilian Vieth-Ditlmann, Fangt einfach an zu messen, SustAI Magazin #3, 2023, <https://sustain.algorithmwatch.org/fangt-einfach-an-zu-messen/>

24 Shaolei Ren, Wasser ist das neue CO₂, SustAI Magazin #3, 2023, <https://sustain.algorithmwatch.org/wasser-ist-das-neue-co2/>

25 Billy Perrigo, OpenAI Used Kenyan Workers on Less Than \$2 Per Hour to Make ChatGPT Less Toxic, Time Magazine, 18. Januar 2023



Die Systeme verbreiten und verstärken Vorurteile und diskriminierende Aussagen und können zu Diskriminierung beitragen

Technische Systeme diskriminieren selbst nicht; es sind immer Menschen, die andere Menschen diskriminieren. Doch je leistungsfähiger ein System ist, desto größer kann der Einfluss auf die mit ihm interagierenden Menschen sein. Es ist bereits problematisch, wenn ein Übersetzungsprogramm aus dem Englischen „doctor“ immer mit Arzt, „nurse“ immer mit Krankenschwester ins Deutsche übersetzt, denn das eine Wort kann auch Ärztin, das andere Krankenpfleger bedeuten. Die Fähigkeiten generativer Systeme reichen viel weiter und können daher auch stärkere Wirkung entfalten, etwa indem Sachverhalte diskriminierend oder stereotyp dargestellt werden. Sehr eindrückliche Beispiele liefern Bildgeneratoren wie MidJourney. Auf die Anweisung (den Prompt) hin, ein Bild zu erzeugen, das ein*e Präsidentschaftskandidat*in zeigt, die*der eine Rede auf einer Straße hält („a presidential candidate is giving a speech in the street“), produzierte das System kein einziges Bild mit einer Frau.²⁶

Die Vertraulichkeit von Daten ist nicht garantiert

Je nach gewünschtem Ergebnis, müssen einer generativen KI mehr oder weniger detaillierte Anweisungen gegeben werden. Soll das System einen Text zusammenfassen, muss dieser Text selbstverständlich auch in vollem Umfang zur Verfügung gestellt werden, ebenso Programmiercode, der auf Fehler geprüft werden soll, oder Daten, um sie visualisieren zu lassen. In derartigen Fällen ist so gut wie nie wirklich klar, ob die zur Verfügung gestellten Informationen auch gespeichert und dazu verwendet werden, die Systeme zu trainieren. Die Nutzungsbedingungen sind fast immer sehr lang und schwer verständlich, oft genug auch mehrdeutig. Einige Unternehmen verbieten deshalb die Nutzung von ChatGPT oder ähnlichen KI-Tools, wie etwa Samsung.

Unklare Rechte an mithilfe generativer KI erzeugten Inhalten

Die Diskussion darüber, welche Inhalte Entwickler*innen von KI-Systemen für das Training ihrer Modelle verwenden durften und dürfen, hat gerade erst begonnen. Dass hier noch vieles unklar ist und auch auf Jahre hinaus bleiben dürfte, führt für die Anwender*innen dazu, dass auch sie mit Rechtsunsicherheiten leben müssen. Sollten etwa Gerichte entscheiden, dass Rechteinhaber*innen, deren Werke für das Training verwendet wurden, auch Rechte an den erzeugten Werken haben, kann das dazu führen, dass Ansprüche auch gegenüber denjenigen entstehen, die sie erzeugen. Nach derzeitigem Stand ist das unwahrscheinlich, sollte aber beobachtet werden.

Geschäfts- und Nutzungsbedingungen

Ein größeres Problem sind die Geschäfts- und Nutzungsbedingungen der Anbieter*innen selbst, also OpenAI, Microsoft, Google, Adobe etc. Wie bei diesen Diensten üblich, sind sie meist sehr komplex und für Laien schwer zu durchschauen, zudem gerade bei kleineren Anbieter*innen nicht auf den jeweiligen Rechtsrahmen angepasst (z.B. werden Bedingungen für die USA formuliert und nur ins Deutsche übersetzt, berücksichtigen aber nicht die in Deutschland geltenden gesetzlichen Bedingungen). Sie sollten daher immer juristisch auf Fallstricke geprüft werden.

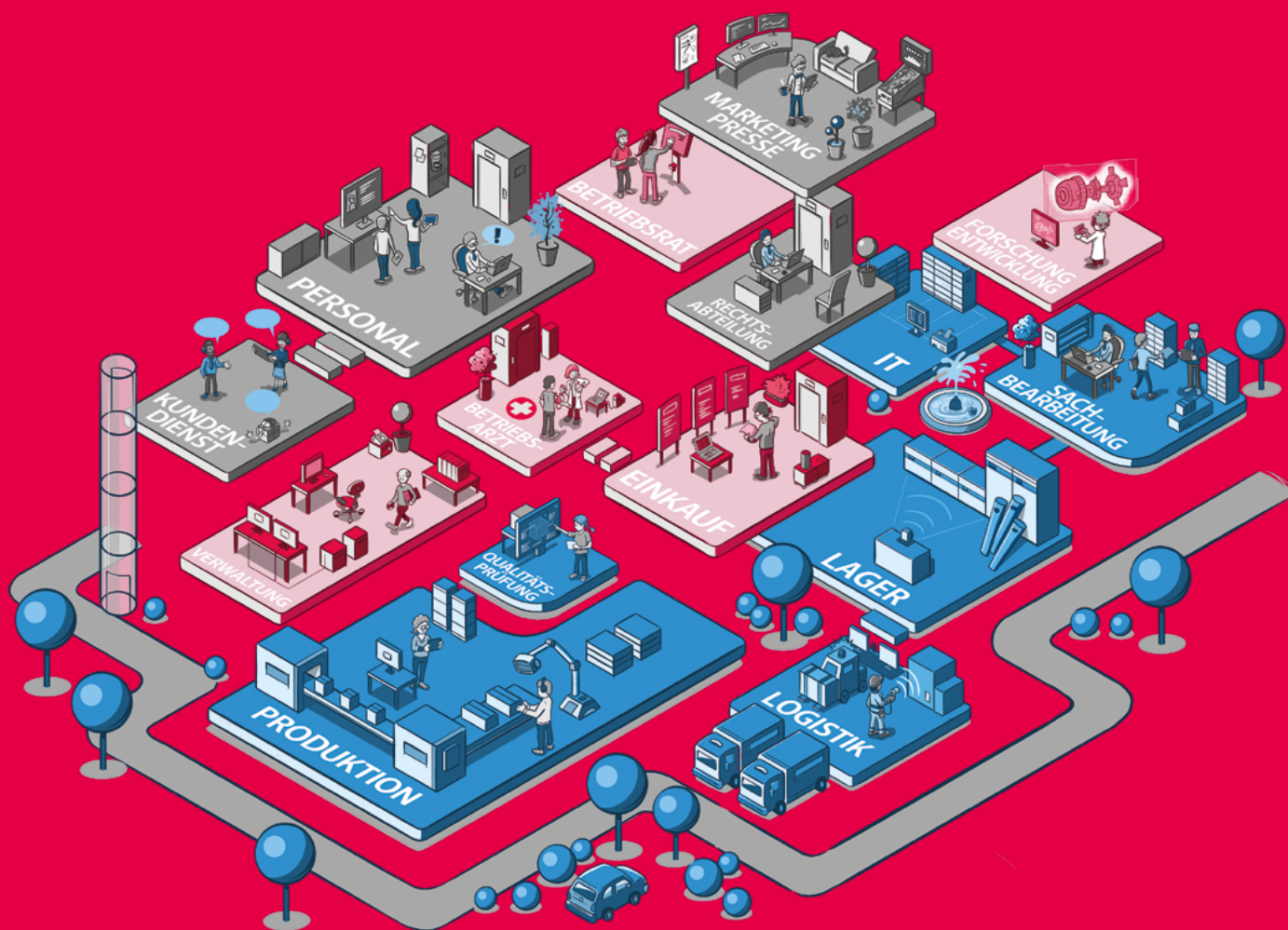
²⁶ Naiara Bellio/Nicolas Kayser-Bril: Some image generators produce more problematic stereotypes than others, but all fail at diversity, 2023, <https://algorithmwatch.org/en/image-generators-stereotypes-diversity/>

Verlust von Arbeitsplätzen

Mit vielen technischen Entwicklungen, die dazu genutzt werden können, Aufgaben zu automatisieren, geht die Sorge einher, dass in der Folge Arbeitsplätze oder auch ganze Berufe verloren gehen. Diese Sorge war auch immer wieder berechtigt; oft haben Technologien große soziale Verwerfungen mit sich gebracht, unter denen viele Menschen gelitten haben. Die Vermutung liegt nahe, dass von generativer KI insbesondere Berufe betroffen sind, in denen der Anteil der kognitiven Arbeit sehr hoch ist. Auch die Annahme, dass ganze Berufsbilder wie etwa das des*der Illustrator*in insgesamt gefährdet sind, erscheint nicht absurd (in dem Sinne, dass nur noch sehr wenige übrig bleiben könnten, die von ihrer Arbeit leben können). Hier soll auf diese sehr komplexen Zusammenhänge allerdings nur hingewiesen werden; angemessen diskutiert werden können sie im Rahmen dieser Publikation nicht.

Für die Risiken **De-Skilling** und **Verlust politischer Handlungsfähigkeit** siehe Kapitel 4.

WELCHE FOLGEN HAT
GENERATIVE KI
FÜR DIE ARBEIT?





WELCHE FOLGEN HAT GENERATIVE KI FÜR DIE ARBEIT?

Veränderung der Arbeit

Die Auswirkungen von KI auf Beschäftigung werden ambivalent diskutiert. Einerseits werden **Risiken** betont, die aufgrund des Automatisierungspotenzials beruflicher Tätigkeiten bestehen und zum Verlust von Arbeitsplätzen führen können. Hierfür kommen insbesondere standardisierbare Routinetätigkeiten, die sich gut in algorithmische Regeln übertragen lassen, in Betracht. Als ein weiteres Risiko wird mit Blick auf die Arbeitsqualität die Gefahr eines „digitalen Taylorismus“²⁷ diskutiert, bei dem die Beschäftigten nur noch nicht automatisierbare „Resttätigkeiten“ verrichten, bei denen keine oder nur begrenzte Autonomiespielräume bestehen und die zudem mit einem erweiterten Überwachungs- und Kontrollpotenzial behaftet sind. Deshalb kann der Einsatz von (generativer) KI auch dazu führen, dass „gläserne Belegschaften“ entstehen – im Übrigen ein Begriff, der in der Debatte um den (Beschäftigten-)Datenschutz bereits seit mehr als 30 Jahren verwendet wird.²⁸ Aufseiten der **Chancen** wird betont, dass KI auch Beschäftigung schaffen könne, weil sie Unternehmen dabei unterstützt, neue Produkte und Dienstleistungen zu kreieren, wodurch zusätzliche Marktpotenziale erschlossen werden können. Dabei könne KI den Menschen produktiver machen, weil mithilfe von Algorithmen besser bzw. schneller gearbeitet werden könne und mit KI somit ein Produktivitätseffekt verbunden sei.²⁹ Neben dieser positiven Wirkung auf die Nachfrageseite am Arbeitsmarkt werden auch Potenziale von KI für die Arbeitsqualität der Beschäftigten diskutiert³⁰, die sich auf das Unterstützungs- und Entlastungspotenzial und damit den „Werkzeugcharakter“ für die Arbeit der Beschäftigten beziehen. Dies schließt auch die Möglichkeit zur Erweiterung des Tätigkeitsspektrums („Augmentation“) und einer Steigerung des Abwechslungsreichtums der Arbeit ein.³¹ Künstliche Intelligenz stellt sich dann vor allem als eine Frage der menschenzentrierten Arbeitsgestaltung.³²

Automatisierung vs. Erweiterung des Tätigkeitsspektrums

In der Diskussion um die Beschäftigungseffekte generativer KI liegt die Gewichtung der Chancen und Risiken der neuen Technologien jedoch etwas stärker auf den Chancen. So werden in einer von der International Labour Organization (ILO) herausgegebenen Studie (Gmyrek et al. 2023) die Potenziale generativer KI als ergänzende Werkzeuge betont, die keineswegs die Folge als „Jobkiller“ haben müssten, sondern auch weitreichende Beschäftigungsmöglichkeiten entfalten können. Für den Einsatz generativer KI eignen sich insbesondere Büroarbeitsplätze. Dort

27 Stefan Kirchner/Sophie-Charlotte Meyer/Anita Tisch, Digitaler Taylorismus für einige, digitale Selbstbestimmung für die anderen? Ungleichheit in der Autonomie in unterschiedlichen Tätigkeitsdomänen, Dortmund 2020.

28 Wolfgang Däubler, Gläserne Belegschaften? Datenschutz für Arbeiter, Angestellte und Beamte, Köln 1987.

29 Peter Buxmann/Holger Schmidt, »Ökonomische Effekte der Künstlichen Intelligenz«, in: dies., Künstliche Intelligenz, Berlin, Heidelberg 2019, S. 21–37.

30 Sophie-Charlotte Meyer/Matthias Hartwig/Anita Tisch/Sascha Wischniewski, »VIII. Künstliche Intelligenz am Arbeitsplatz: Verbreitung und Hinweise auf Zusammenhänge mit Arbeitsqualität«, in: Anita Tisch/Sascha Wischniewski (Hg.), Sicherheit und Gesundheit in der digitalisierten Arbeitswelt. Kriterien für eine menschengerechte Gestaltung, Baden-Baden 2022, S. 313–334.

31 Martina Frost/Kerstin Guhle/Anja Cordes/Katrin Zittlau/Oliver Hasselmann, »Produktive, sichere und gesunde Arbeitsgestaltung mit digitalen Technologien und Künstlicher Intelligenz – Hintergrundwissen und Gestaltungsempfehlungen« in: Zeitschrift für Arbeitswissenschaft 74, Nr. 2 (2020), S. 76–88.

32 Lars Adolph/Alina Tausch, »Künstliche Intelligenz in der Arbeitswelt«, in: Eva Bamberg/Antje Ducki/Monique Janneck (Hg.), Digitale Arbeit gestalten. Herausforderungen der Digitalisierung für die Gestaltung gesunder Arbeit, Wiesbaden 2022, S. 33–48; Walter Ganz/David Kremer/Markus Hoppe/Anne-Sophie Tombeil/Claudia Dukino/Helmut Zaiser/Claus Zanker, Arbeits- und Prozessgestaltung für KI-Anwendungen, Stuttgart 2021.

sei etwa ein Viertel der Arbeitsplätze betroffen – in die eine oder andere Richtung. Der Studie zufolge dürften insbesondere assistierende Bereiche in der Sachbearbeitung zugänglich für generative KI-Anwendungen sein. Gezeigt wird, dass insbesondere bei assistierenden Bereichen in der Sachbearbeitung, für die keine oder nur eine geringe berufliche Qualifikation erforderlich ist, die Automatisierungseffekte überwiegen dürften, während für das obere Qualifikationssegment erwartet wird, dass sich der Einsatz generativer KI stärker im Augmentationsszenario bewegt, also das Tätigkeitsspektrum sich durch die Nutzung generativer KI eher erweitert. In nachfolgender Abbildung sind berufliche Tätigkeiten dargestellt, für die in der Studie angenommen wird, dass generative KI eher zur Automatisierung von Arbeitsaufgaben beiträgt:

Abbildung 2: Tätigkeitsgruppen mit hohem Automatisierungspotenzial beim Einsatz generativer KI

Tätigkeitsgruppe	Beispieltätigkeiten
Administrative und kommunikative Tätigkeiten	▪ Terminvereinbarungen mit Kund*innen ▪ Selbständige Erledigung von Routinekorrespondenz ▪ Kopieren von Dokumenten ▪ Adressierung von Rundschreiben und Umschlägen von Hand
Datenmanagement und Buchhaltung	▪ Buchführung über Lagerbestände und finanzielle Transaktionen ▪ Anlegen von Akten bei Neueinstellungen; Überprüfung von Akten auf Vollständigkeit ▪ Datenimport und-export zwischen verschiedenen Datenbanken und Software ▪ Erstellung von Rechnungen und Kaufverträgen sowie Annahme von Zahlungen
Informationsverarbeitung und Übersetzung	▪ Aufnahme von Diktaten und stenografischen Aufzeichnungen ▪ Übersetzung von einer Sprache in eine andere ▪ Umwandlung von Informationen in Codes und Klassifizierung von Informationen nach Codes für Datenverarbeitungszwecke ▪ Eingeben von Anweisungen zur Programmierung elektronischer Geräte ▪ Erfassen, Aufbereiten, Sortieren, Klassifizieren und Ablegen von Informationen
Bereitstellung von Informationen und Beantwortung von Anfragen	▪ Beantwortung von Anfragen zu Problemen; Bereitstellung von Ratschlägen, Informationen und Unterstützung ▪ Ermittlung des Kundenbedarfs; Beratung zur Produktpalette, Preis, Lieferung, Garantien sowie Produktnutzung und-pflege ▪ Beantwortung von Anfragen zu Dienstleistungen und Kosten für Raum- und Gerätemiete, Catering und ähnlichen Dienstleistungen

Quelle: adaptiert nach Gmyrek/Berg/Bescond 2023³³

33 Pawel Gmyrek/Janine Berg/David Bescond, Generative AI and jobs: A global analysis of potential effects on job quantity and quality, Geneva 2023, S. 22f.



Welche Folgen hat generative KI für die Arbeit?

Während der Einsatz generativer KI auf der einen Seite zur Automatisierung von Tätigkeiten führen kann, lassen entsprechende Anwendungen auf der anderen Seite eine Erweiterung bzw. Augmentation des Tätigkeitsspektrums von Beschäftigten erwarten. In der Studie werden insbesondere für die nachfolgend genannten Berufe bzw. Berufsgruppen Augmentationspotenziale prognostiziert:

Abbildung 3: Tätigkeitsgruppen mit hohem Erweiterungspotenzial beim Einsatz generativer KI

Tätigkeitsgruppe	Beispieltätigkeiten
Büroangestellte	▪ Lagerist*innen, Lagerverwalter*innen
Manager*innen	▪ Führungskräfte in den Bereichen Versorgung, Vertrieb und Ähnliches
Fachkräfte	▪ Berufsbildung, Lehrer*innen ▪ Software- und Anwendungsentwickler*innen, Analyst*innen ▪ Computernetzwerk-Fachleute ▪ Fachleute für Bildungs- und Erziehungsmethoden
Einfache Tätigkeiten	▪ Ableser*innen von Zählerständen, Automatenaufsteller*innen ▪ Straßenarbeiter*innen und verwandte Berufe ▪ Bot*innen, Kuriere und Gepäckträger*innen
Anlagen- und Maschinenbediener*innen, Monteur*innen	▪ Elektromonteur*innen, Elektroniker*innen ▪ Gabelstaplerfahrer*innen ▪ Maschinenbediener*innen für fotografische Produkte ▪ Bediener*innen für mechanische Maschinen
Service- und Verkaufsmitarbeiter*innen	▪ Verkäufer*innen in Supermärkten und an Verkaufsständen ▪ Transport- und Verkehrsleiter*innen ▪ Gebäudebetreuer*innen
Techniker*innen und vergleichbare Berufe	▪ Rechts- und Verwaltungsberufe ▪ Rundfunk- und Fernsehtechniker*innen ▪ Telekommunikationstechniker*innen ▪ Produktionsleiter*innen ▪ Beschäftigte des öffentlichen Gesundheitsdienstes

Quelle: adaptiert nach Gmyrek/Berg/Bescond 2023³⁴

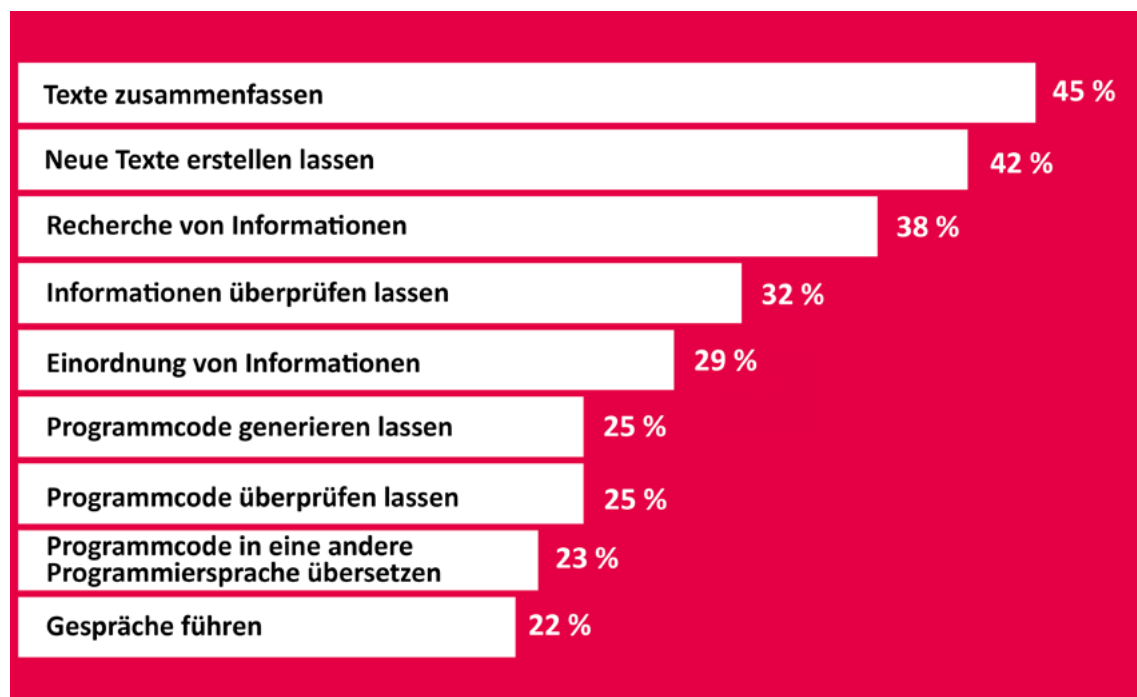
34 Pawel Gmyrek/Janine Berg/David Bescond, Generative AI and jobs: A global analysis of potential effects on job quantity and quality, aaO. (FN 32), S. 29.

KI als Unterstützung

Die unterschiedlichen Wirkrichtungen, die generative KI auf Arbeit und Beschäftigte haben kann, verdeutlichen sich auch in Studien, die empirisch auf Basis subjektiver Problemwahrnehmungen von Beschäftigten argumentieren. So wird in einer Studie des Bayerischen Forschungsinstituts für Digitale Transformation³⁵ gezeigt, dass rund ein Viertel aller Erwerbstätigen bereits generative KI im beruflichen Kontext genutzt hat. Es zeigt sich, dass persönliche Erfahrungen im Umgang mit generativer KI das Problembewusstsein für die Potenziale, aber auch die möglichen Folgen für die eigene Tätigkeit schärft. Demnach erwarten Beschäftigte, die bereits textbasierte generative KI im Beruf genutzt haben, überdurchschnittlich häufig, dass ihre beruflichen Tätigkeiten mindestens teilweise ersetzt werden (77 %). Gleichwohl wird die Nutzung generativer KI im beruflichen Kontext der Studie zufolge von der Mehrheit der Nutzer*innen als gewinnbringend wahrgenommen. So stimmen 60 Prozent der Nutzer*innen der Aussage zu, dass sie die von generativer KI erzeugten Ergebnisse sinnvoll im Beruf nutzen konnten. 64 Prozent konnten dadurch Arbeitszeit sparen. Mehr als die Hälfte (56 %) erlebt generative KI als Unterstützung bei Aufgaben, die ihnen schwerfallen. 63 Prozent stimmen der Aussage voll und ganz oder teilweise zu, dass die mit Unterstützung generativer KI produzierten Ergebnisse positiver wahrgenommen wurden.

Abbildung 4: Beruflicher Nutzungszweck textbasierter generativer KI bei Erwerbstätigen

Frage: Bitte geben Sie an, in welchem Bereich und für welche Zwecke Sie „generative KI“-Systeme für die Erstellung und/oder Prüfung von Texten oder Programmcode verwendet haben. Mehrfachnennung möglich – rein berufliche Nutzung



Quelle: Bayerisches Forschungsinstitut für digitale Transformation [bidit]³⁶

35 Antonia Schlude/Mara Schwind/Ulrike Mendel/Roland A. Stürz/Danilo Harles/Micha Fischer, Verbreitung und Akzeptanz generativer KI in Deutschland und an deutschen Arbeitsplätzen, <https://www.bidit.digital/publikation/verbreitung-und-akzeptanz-generativer-ki-in-deutschland-und-an-deutschen-arbeitsplaetzen/> (Zugriff am 26.03.2024).

36 Antonia Schlude/Mara Schwind/Ulrike Mendel/Roland A. Stürz/Danilo Harles/Micha Fischer, Verbreitung und Akzeptanz generativer KI in Deutschland und an deutschen Arbeitsplätzen, <https://www.bidit.digital/publikation/verbreitung-und-akzeptanz-generativer-ki-in-deutschland-und-an-deutschen-arbeitsplaetzen/>, (Zugriff am 26.03.2024).



Markteffekte von KI

Bislang gibt es nur Mutmaßungen darüber, wie sich generative KI in einzelnen Branchen und Berufsfeldern auswirken wird. Die potenziellen transformativen Effekte generativer KI für Branchen und Berufsgruppen war jedoch gelegentlich Gegenstand wissenschaftlicher Untersuchungen. So wird etwa erwartet, dass generative KI die Nachfrage nach bestimmten Dienstleistungen, wie sie nicht selten von Solo-Selbständigen erbracht werden, nachhaltig negativ beeinflussen kann. Zu solchen Diensten zählen etwa Schreibarbeiten, die Software-, App- und Webentwicklung, Ingenieurarbeit oder auch das Grafikdesign und die 3D-Modellierung³⁷. Die Auswertung erfolgte anhand von Daten zu über einer Million Aufträgen, die zwischen Juli 2021 und Juli 2023 auf einer führenden Online Labour Market-Plattform ausgeschrieben wurden. Im Ergebnis zeigt sich, dass die durchschnittlichen Rückgänge der wöchentlichen Auftragszahlen infolge der Nutzung generativer KI-Tools bei Schreibarbeiten 30,4 Prozent, bei der Software-, App- und Webentwicklung 20,6 Prozent, bei der Ingenieursarbeit 10,4 Prozent, im Grafikdesign 18,5 Prozent und bei der 3D-Modellierung 15,5 Prozent betrugen. Jedoch werden in branchenbezogenen Analysen nicht nur die möglichen Substitutionseffekte betrachtet. So konnte in einer Untersuchung des Fraunhofer-Instituts für Arbeitswirtschaft und Organisation gezeigt werden, dass generative KI hohe Unterstützungspotenziale für das Innovationsmanagement im Hinblick auf analytisch und/oder kreativ geprägte Aufgaben haben kann.³⁸ Generative KI im Innovationsmanagement kann etwa dabei helfen, Probleme zu identifizieren und auszuwählen, bestehende Lösungsoptionen zu identifizieren, neue Lösungsoptionen zu generieren oder Lösungskonzepte zu erstellen und zu testen.

Frauen sind von KI in besonderer Weise betroffen

Aus der vorliegenden Studienlage geht zudem hervor, dass insbesondere die Arbeitsplätze von Frauen Substitutionsrisiken ausgesetzt sind, da die zur Automatisierung geeigneten Tätigkeiten insbesondere in Berufen vorkommen, die überdurchschnittlich häufig von Frauen ausgeführt werden. So wird anhand von Mikrodaten der ILO gezeigt, dass 3,7 Prozent der weltweit von Frauen ausgeführten Beschäftigung potenziell mit generativer KI automatisierbar ist, während dies bezogen auf männliche Beschäftigte lediglich auf 1,4 Prozent zutrifft. Gleichzeitig ist auch der Anteil an Jobs mit hohem Augmentationspotenzial bei Frauen größer als bei Männern. Daraus wird in der Studie abgeleitet, dass jegliche Form technologischen Wandels in hohem Maß von Geschlechtereffekten begleitet ist. Dies legt nahe, dass Gestaltungsdefizite unverhältnismäßig große Nachteile für weibliche Beschäftigte nach sich ziehen, während aus einer gelungenen Gestaltung des Wandels Potenziale für eine Stärkung der Rolle von Frauen im Arbeitsprozess verbunden sein können.³⁹ KI bietet jedenfalls einen Anlass, „Geschlechterungleichheiten neu zu verhandeln.“⁴⁰ Dies kann einerseits in Form normativer Leitbilder von „richtiger“ oder „guter“ Arbeit erfolgen, indem beispielsweise KI-basierte Assistenzsysteme weiblich vergeschlechtlicht konnotiert werden („Alexa“, „Siri“). Auf der anderen Seite kann KI auch diskriminierend wirken und Geschlechterstereotype reproduzieren, weil sie häufig mit Daten trainiert werden, in die kulturelle Muster eingeschrieben sind.

37 Jonas Hannane/Ozge Demirci/Xinrong Zhu, Generative Künstliche Intelligenz reduziert Nachfrage nach Freelance-Arbeit auf Online-Plattformen, Berlin 2024.

38 Lukas Keicher/Antonio Ardillio/Markus Korell, Generative KI im Innovationsmanagement – Potenziale und Herausforderungen, Stuttgart 2024.

39 Pawel Gmyrek/Janine Berg/David Bescond, Generative AI and jobs: A global analysis of potential effects on job quantity and quality, aaO. (FN 32), S. 34f.

40 Tanja Carstensen/Kathrin Ganz, »Künstliche Intelligenz und Gender – eine Frage diskursiver (Gegen-)Macht?« in: WSI-Mitteilungen 77, Nr. 1 (2024), <https://www.nomos-elibrary.de/10.5771/0342-300X-2024-1-26.pdf>, (Zugriff am 07.02.2025), S. 26–33, 29.

Machtverschiebung durch KI

Doch welche Konsequenzen haben die zunehmenden Fähigkeiten (generativer) KI-Systeme für das Verhältnis von Menschen und Technik im Arbeitsprozess? „Übernimmt jetzt die KI?“, wie als Titel von Tagungen⁴¹ oder Podcasts⁴² gelegentlich provokant gefragt wurde. Werden die arbeitenden Menschen tatsächlich von immer leistungsfähigeren Anwendungen aus den Betrieben verdrängt oder zu Handlangern der KI-Systeme degradiert? Antworten auf diese Fragen sind auch für das Handeln von Betriebs- und Personalrät*innen entscheidend, weil sie darüber bestimmen, wie sie sich der betrieblichen Nutzung generativer KI-Systeme gegenüber positionieren und verhalten. Für die perspektivische Beschäftigungssituation im Betrieb hilft ein Blick auf machttheoretische Überlegungen aus der Arbeitssoziologie. Dort wird die These formuliert, dass es durch den Einsatz generativer KI zu einer Machtverschiebung bei der Zusammenarbeit von Menschen und Technik kommen kann.⁴³ Dabei sind zwei Fragen zentral: Inwieweit findet eine Machtverschiebung statt, indem einer KI-Einfluss zugeschrieben oder gewährt wird? Kann generative KI die für den Menschen relevanten Unsicherheitszonen besetzen?

Einfluss von KI auf Sphären des Arbeitshandelns

Mit Bezug auf die erste Frage lässt sich die Übertragung von Einfluss bzw. „Macht“ auf bestimmte Tätigkeitsfelder mit der Leistungsfähigkeit generativer KI begründen. Als Beispiel wird das Sprachverständnis angeführt, bei dem KI inzwischen die menschlichen Fähigkeiten weit überholt hätte. Gleichwohl verfügen generative KI-Systeme aufgrund eines fehlenden Bewusstseins über kein Eigeninteresse, selbst Macht ausüben zu wollen. Die Übertragung von Macht und Einfluss auf die KI erfolgt somit im Eigeninteresse des Menschen, also im Vertrauen in die größere Leistungsfähigkeit generativer KI-Systeme. „Der Mensch gesteht einer KI beziehungsweise den Agenten, die die KI einsetzen, Einflussmacht zu, wenn sich seine Nutzenposition durch die Akzeptanz der Einflussmacht verbessert. Insofern kann es ökonomisch effizient sein, eigene Entscheidungsräume einzugrenzen, um dadurch Komplexität zu reduzieren.“⁴⁴

41 ver.di, FilmUnion-Wissen & Austausch: Übernimmt jetzt die KI? <https://medien-kultur-bb.verdi.de/service/seminare-termine/++co++3a60e3f6-7434-11ef-a878-33e04bf09613> (Zugriff am 07.02.2025); BLM, Übernimmt jetzt die KI? Auswirkungen der Künstlichen Intelligenz auf Medien und Gesellschaft, https://www.blm.de/files/pdf2/einladung_final_low.pdf (Zugriff am 07.02.2025).

42 Valerie Wagner/Heike Stiegler, Übernimmt jetzt die KI? Format follows Story Podcast, Folge 17, <https://open.spotify.com/episode/5mo514g1OOWmB7gyKM8oaa> (Zugriff am 07.02.2025).

43 Ulrich Lenz, »Machtverschiebung durch generative künstliche Intelligenz- und die Konsequenzen für das Transformationsmanagement«, in: Olaf Geramanis/Stefan Hutmacher/Lukas Walser (Hg.), Organisationale Machtbeziehungen im Wandel. Führung zwischen Zustimmung und Zwang, Wiesbaden, Heidelberg 2023, S. 291–307.

44 Ulrich Lenz, Machtverschiebung durch generative künstliche Intelligenz- und die Konsequenzen für das Transformationsmanagement, aaO. (FN 42), S. 296.



Kontrolle der Handlungs- und Entscheidungsautonomie des arbeitenden Menschen

In der zweiten Frage wurde die Besetzung relevanter Unsicherheitszonen angesprochen. Deren Kontrolle gilt in der Machtkonzeption von Michel Crozier und Erhard Friedberg⁴⁵ als die zentrale Quelle von Macht in Organisationen. Als Unsicherheitszonen werden dabei die spezifischen Handlungsspielräume verstanden, über die alle Mitglieder einer Organisation verfügen. Diese für andere nicht kontrollierbaren Zonen der Unsicherheit bzw. Ungewissheit gelten dabei als Machtressource derjenigen, die sie kontrollieren. Nach Crozier und Friedberg kann sich Macht aus unterschiedlichen Quellen speisen: „(1) aus *Expertentum*, das aus der Beherrschung eines spezifischen Sach- oder Fachwissens und der funktionalen Spezialisierung herrührt, (2) aus *Nahtstellen zur Umwelt*, die von einzelnen Organisationsmitgliedern besetzt werden, und deren Bedeutung für die Organisation, (3) aus der *Kontrolle von Informations- und Kommunikationskanälen* in der Organisation und (4) aus den *Organisationsregeln*.“⁴⁶ Die Ausübung von Macht bedeutet dann, Einfluss auf die von anderen kontrollierten Unsicherheitszonen zu gewinnen.

Übertragen auf das Eindringen generativer KI in vom Menschen kontrollierte Unsicherheitszonen (z.B. Wissensdomänen) würde eine Machtausweitung bzw.-übernahme von KI bedeuten, Outputs zu generieren, die den Menschen zur Verhaltensanpassung veranlassen würde. In der Praxis werden bereits erste Anzeichen erkannt, dass generative KI das Potenzial hat, menschlich relevante Unsicherheitszonen zu besetzen. Beispielhaft wird auf die Fähigkeit von Large Language Models (LLM) verwiesen, Anwendungsprogramme vollständig automatisiert zu erstellen und zu programmieren. „Da das LLM auch die Fähigkeit hat, einen bestehenden Programmcode zu analysieren, kann man hier schon von einer Besetzung menschlicher Unsicherheitszonen sprechen, sobald sich die Qualität des Programmierens weiter erhöht. Die generative KI übernimmt dann nicht nur das Programmieren, sondern auch die automatisierte Analyse von Fehlern bei bestehenden Programmcodes. Die KI kann dann den Menschen anweisen, an welchen Stellen eine Programmpformance verbessert werden kann.“⁴⁷ In der Debatte um die Machtverschiebung aufgrund des Einsatzes generativer KI stimmt jedoch die zugrunde gelegte relationale Machtkonzeption von Crozier und Friedberg hoffnungsvoll, da Machtbeziehungen – im Betrieb oder außerhalb – niemals einseitig verstanden werden können, sondern immer als Tausch gelten und auch für diejenigen, auf die Macht ausgeübt wird, einen Mehrwert stiften müssen.

Wenn sich Menschen KI-Entscheidungen unterordnen, tun sie dies dem theoretischen Konzept folgend nur, weil sie sich davon einen Vorteil versprechen. Diese Erkenntnis ist nicht neu. Bezogen auf betriebliche Informationssysteme und deren Eignung als Instrument betrieblicher Machtpolitik stellte der Organisationssoziologe Günter Ortmann bereits 1987 fest: „Es gibt gute Gründe für die Vermutung, daß Orwellsche Kontrollbedürfnisse eher aus Angst vor dem Chaos denn aus ökonomischer Vernunft Platz greifen, daß sie in den Diskussionen der letzten Jahre stark überschätzt worden sind, daß sie relativ gut eindämmbar sind, und daß Management- und Betriebsratsinteressen – durchaus in weiten Teilen übereinstimmend – auf Bereinigung von Unsicherheitszonen, auf wohlgeordnete und-geplante Produktion, auf die Vermeidung der Inanspruchnahme von Drohpotentialen und auf eine sichere Prozeßbeherrschung durch das Management zielen.“⁴⁸ Diese damals zutreffende Feststellung lässt sich auch auf den betrieblichen Einsatz generativer KI-Systeme übertragen und mahnt – bei aller gebotenen Vorsicht im Hinblick auf die diskutierten Risiken für Beschäftigte – zu einer gewissen Gelassenheit im Umgang mit ihnen an.

45 Michel Crozier/Erhard Friedberg, *Macht und Organisation. Die Zwänge kollektiven Handelns*, Königstein/Ts. 1979.

46 Alfred Kieser/Peter Walgenbach, *Organisation*, Stuttgart 2010, S. 53, H.i.O.

47 Ulrich Lenz, *Machtverschiebung durch generative künstliche Intelligenz- und die Konsequenzen für das Transformationsmanagement*, aaO. (FN 42), S. 301.

48 Günther Ortmann, „Betriebliche Informationssysteme und Machtpolitik im Betrieb“, in: Jürgen Friedrichs (Hg.), 23. Deutscher Soziologentag „Technik und sozialer Wandel“. Hamburg, 1986: Sektions- und Ad-hoc-Gruppen, Opladen 1987, S. 430–431.

Auswirkungen auf die Selbstbestimmtheit des arbeitenden Menschen

Schon sehr lange wird über die Frage gestritten, ob Werkzeuge im Allgemeinen und Automatisierung im Besonderen zum so genannten De-Skilling der Menschen beitragen, sie also am Ende darauf reduzieren werden, triviale und fremdbestimmte Arbeit zu verrichten, oder sogar als Arbeitskraft vollständig überflüssig werden lassen – mit all den daraus folgenden Verwerfungen für die Einzelnen und die Gesellschaft. Jedes Mal, wenn ein bedeutendes neues Werkzeug entwickelt wird, wird die Frage neu gestellt und mit Verve diskutiert, so auch im Angesicht generativer KI. Die Frage ist weder überflüssig noch einfach – und viel weniger noch eindeutig – zu beantworten, denn die Lehren der Vergangenheit sind zwar hilfreich und sollten nicht vergessen werden, bestimmen aber nicht den Verlauf der Zukunft.

In diesem Kapitel werden wir die Diskussionen darüber, ob neue Werkzeuge dazu beigetragen, dass uns Menschen die Arbeit ausgeht und ob KI die „Weltherrschaft“ übernehmen wird, ausklammern. Unsere Antwort ist: Weder das eine noch das andere. Das zu begründen, würde sehr viel Raum einnehmen. Wir verweisen daher auf die ausgezeichnete Darstellung der Positionen in Michael Seemanns Working Paper der Forschungsförderung der Hans-Böckler-Stiftung mit dem Titel „Künstliche Intelligenz, Large Language Models, ChatGPT und die Arbeitswelt der Zukunft“.⁴⁹

Stattdessen werden wir uns auf folgende Fragen fokussieren:

- Kann und wird generative KI zu einem mehr oder zu einem weniger selbstbestimmten Arbeiten beitragen? (De-Skilling)
- Welche Auswirkungen kann und wird generative KI darauf haben, ob Unternehmen und Gesellschaft über die Zukunft (mit-)bestimmen können?

⁴⁹ Michael Seemann, Künstliche Intelligenz, Large Language Models, ChatGPT und die Arbeitswelt der Zukunft, Working Paper Forschungsförderung 304, Hans-Böckler-Stiftung, Düsseldorf 2023, insbesondere Abschnitte 3.2 und 4.2.



De-Skilling oder Augmentierung

Der Begriff De-Skilling wird meist verwendet, um den Prozess zu beschreiben, in dem qualifizierte durch unqualifizierte Arbeitskräfte ersetzt werden. Hier geht es jedoch um den Prozess, in dem Menschen „infolge eines immer durchgreifenderen Einsatzes von KI-Anwendungen in unterschiedlichen Lebensbereichen [...] mehr und mehr dazu verleitet werden könnten, eigene Aufgaben an die Technik zu delegieren, weil diese als (vermeintlich) überlegen angesehen wird,“ wie es der Deutsche Ethikrat beschreibt. Hieraus ergebe sich „ein Risiko für die individuelle Selbstwahrnehmung. Mit zunehmender Delegation von Aufgaben an Technologien kann der Eindruck entstehen, immer mehr Kontrolle über wesentliche Bereiche des eigenen Lebens abzugeben.“⁵⁰

Das kann auch auf den betrieblichen Kontext übertragen werden. Der Ethikrat weist richtigerweise darauf hin, dass ein solcher Kompetenzverlust „nicht allein auf digitale oder gar KI-basierte Technologien begrenzt“ ist, sondern bei nahezu jedem Einsatz von Werkzeugen beobachtet werden kann.

Dennoch soll die Empfehlung zum Thema in voller Länge zitiert werden: „Ob und inwiefern beim Einsatz von KI-Anwendungen Verluste menschlicher Kompetenz auftreten, die als unerwünscht eingestuft werden, muss sorgfältig beobachtet werden. Bei der Entwicklung und dem Einsatz neuer Technologien sind solch unerwünschte Kompetenzverluste durch eine sinnvolle Gestaltung des Zusammenspiels von Mensch und Technik, durch angemessene institutionelle und organisatorische Rahmenbedingungen sowie durch gezielte Gegenmaßnahmen wie etwa spezifische Trainingsprogramme zu minimieren bzw. zu kompensieren. Kompetenzverluste können sowohl individueller als auch kollektiver Natur sein. So gilt es zu verhindern, dass die Delegation von Aufgaben an Technologien dazu führt, dass Gesellschaften übermäßig anfällig werden, wenn diese Technologien (zeitweise) ausfallen. Jenseits dieser systemischen Aspekte müssen negative Auswirkungen solcher Delegation auf die individuelle Autonomie oder Selbstwahrnehmung mitigiert werden.“⁵¹

Zu den Gestaltungsmöglichkeiten und -aufgaben gehört es, Ideen dazu zu entwickeln, wie Systeme generativer KI so in den Arbeitsalltag integriert werden können, dass sowohl Unternehmen als auch Beschäftigte davon profitieren. Dafür sollten die Voraussetzungen gelten, die auch für den Einsatz anderer KI-basierter Systeme entwickelt wurden:⁵²

- ☐ Beschäftigte haben einen Anspruch darauf, die Logik derartiger Systeme verstehen zu können. Denn das ist eine notwendige – wenn auch keine hinreichende – Voraussetzung dafür, dass die Autonomie der Beschäftigten gewahrt bleibt; sie dürfen keine passiven Objekte einer von Algorithmen getriebenen Steuerung („algorithmic governance“) werden. Vielmehr müssen sie aktiv dazu beitragen können, die Leistung der Organisation zu verbessern, indem sie Erkenntnisse nutzen, die durch den Einsatz dieser algorithmischen Systeme entstehen.
- ☐ Organisationen, die KI-Tools entwickeln und implementieren wollen, müssen Expert*innen mit der Bandbreite an Fähigkeiten (und informellem Wissen und kognitiver Haltung) rekrutieren, die erforderlich sind, um die Verständlichkeit und Fairness der Systeme zu bewerten.
- ☐ Die Auswirkungen der Anwendung des KI-Werkzeugs auf die Mitarbeitenden sollten sorgfältig überwacht werden.⁵³

50 Deutscher Ethikrat, Mensch und Maschine – Herausforderungen durch Künstliche Intelligenz, Stellungnahme, Berlin 2023, S. 356.

51 Deutscher Ethikrat, Mensch und Maschine – Herausforderungen durch Künstliche Intelligenz, Stellungnahme, Berlin 2023, S. 357.

52 AlgorithmWatch: Positionen zum Einsatz von KI im Personalmanagement. Rechte und Autonomie von Beschäftigten stärken – Warum Gesetzgeber, Unternehmen und Betriebsräte handeln müssen, https://algorithmwatch.org/de/wp-content/uploads/2020/03/AlgorithmWatch_AutoHR_Positionspapier_2020.pdf

53 Für Details, s. Michele Loi, People Analytics muss den Menschen zugutekommen. Eine ethische Analyse datengesteuerter algorithmischer Systeme im Personalmanagement, Study 450, Hans-Böckler-Stiftung, Düsseldorf 2021.

Handlungsfähigkeit in Unternehmen und Gesellschaft

Die Empfehlung des Ethikrats verweist auf einen weiteren wichtigen Aspekt der Entwicklung von Technologien, die eine Art „Infrastrukturen“ darstellen. Schon jetzt bestehen Abhängigkeiten von Produkten, die heute in erster Linie als Software as a service (SAAS) angeboten werden: sowohl auf Ebene des Betriebssystems (Microsoft Windows, Apple iOS, Google Android etc.), der Anwendungen (Microsoft, SAP, Oracle, Salesforce etc.) als auch der Dienste (Cloud-Services von Amazon, Google, Microsoft u.a.). Generative KI hat Eigenschaften, die ebenfalls dazu führen können (und sehr wahrscheinlich werden), dass sich eine Oligopolstruktur bei Anbieter*innen herausbilden wird: Generative KI zu entwickeln, ist enorm teuer und setzt die Mitarbeit hochqualifizierter Entwickler*innen voraus. Zusammen mit der Erwartung, dass im Erfolgsfall gewaltige Gewinne zu erwarten sind, führt das zu einem so genannten „Winner takes all“-Markt, in dem am Ende sehr wenige Anbieter*innen übrigbleiben, die den Großteil der Nachfrage bedienen. Das wiederum kann zu Lock-in-Effekten führen, in denen Kosten, zu einem anderen Angebot zu wechseln, extrem hoch sind.

Für Unternehmen besteht derzeit noch die Möglichkeit, Produkte verschiedener Anbieter zu nutzen, auch wenn es wenige sind. Hier gilt es, Erfahrungen mit diesen Angeboten zu sammeln. Zudem sollten Unternehmen, in denen der technische Sachverstand vorhanden ist, Modelle testen, die unter Lizenzen zur Verfügung gestellt werden, die den eigenen Betrieb ermöglichen. Dort, wo der vorhergesehene Nutzen so groß ist, dass der Aufwand als lohnenswert eingeschätzt wird, können auch externe Dienstleister*innen damit beauftragt werden, derartige Varianten umzusetzen – etwa den Betrieb eines generativen Modells auf einem Server unter eigener Kontrolle.

Herausforderung für die betriebliche Mitbestimmung

Wie wird generative Künstliche Intelligenz im betrieblichen Alltag bereits genutzt, welche Herausforderungen gehen mit ihr für betriebliche Interessenvertreter*innen einher und welche Ansätze zur Regulierung und Gestaltung kristallisieren sich als angemessen heraus? Diesen Fragen geht eine im Sommer 2024 von der INPUT Consulting gGmbH durchgeführte Onlinebefragung von Betriebs- und Personalräten nach. An der Befragung, die über den E-Mail-Verteiler des ver.di-Bereichs Mitbestimmung beworben wurde, nahmen insgesamt 180 Personen aus unterschiedlichen Dienstleistungsbranchen teil.

Von unseren Befragten waren insgesamt 96 Prozent in einem Betriebs- oder Personalrat aktiv, 55 Prozent vollständig freigestellt, 13 Prozent waren teilweise freigestellt und 28 Prozent kommen ihrer Gremientätigkeit neben ihrer regulären Beschäftigung nach. Etwa zwei Drittel unserer Befragten sind männlichen Geschlechts, ein Drittel weiblich und ein Prozent definiert sich als divers. Damit liegt der Frauenanteil unter unseren Befragten deutlich unterhalb der durchschnittlichen Frauenquote im Organisationsbereich der Gewerkschaft ver.di, die für die Betriebsratswahlen 2022 mit 47 Prozent angegeben wird (Quelle siehe kommentierte Grundauszählung). Unsere Befragungsteilnehmer*innen stammen überwiegend aus mittelgroßen und großen Unternehmen bzw. Verwaltungen. Etwa die Hälfte gibt an, dass die Betriebsgröße bei 1.000 oder mehr Beschäftigten liegt. Der Anteil an Befragten, die in kleinen und mittleren Betrieben (KMU) tätig sind, ist mit unter zehn Prozent gering. Bezogen auf die Branchenzugehörigkeit ist der öffentliche Dienst mit einem Anteil von 39 Prozent unserer Befragten am stärksten vertreten, gefolgt von den Finanz- und Versicherungsdienstleistungen (13 %), dem Gesundheits- und Sozialwesen (11 %) und dem Handel (8 %). Die übrigen Dienstleistungsbranchen sind mit geringeren Anteilen repräsentiert.

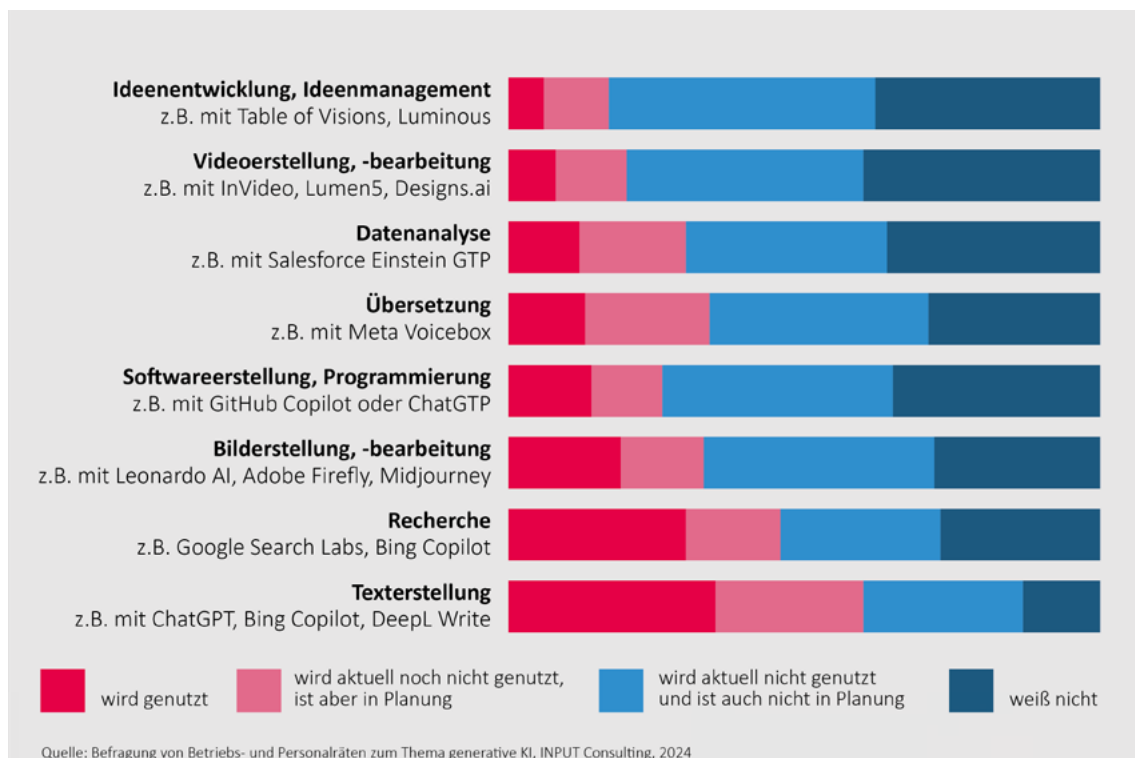


Inwieweit wird generative KI bereits in Betrieben und Verwaltungen verwendet oder existieren Planungen dazu?

Unsere Befragung von betrieblichen Interessenvertreter*innen zeigt, dass auch generative KI inzwischen in Unternehmen und Verwaltungen angekommen ist. Insbesondere sind Systeme, die zur Texterstellung geeignet sind (z.B. ChatGPT, Bing Copilot, DeepL Write), in der Praxis schon weit verbreitet. So geben 35 Prozent der Befragten an, dass entsprechende Anwendungen bereits Teil der betrieblichen Praxis sind, weitere 25 Prozent antworten, dass der Einsatz solcher generativen KI-Systeme für die Zukunft geplant ist (siehe Abbildung 5). Kaum weniger verbreitet ist die Nutzung generativer KI zu Recherchezwecken. Bezogen auf diesen Nutzungszweck geben 31 Prozent der Befragten an, dass Anwendungen wie z.B. Google Search Labs oder Bing Copilot zur Beschaffung von Informationen bereits genutzt werden. Weitere 16 Prozent berichten von Nutzungsabsichten zur Informationsrecherche für die Zukunft. Etwas weniger verbreitet sind Nutzungserfahrungen mit generativen KI-Systemen zur Bilderstellung und -verarbeitung, zur Softwareerstellung und Programmierung, zu Übersetzungszwecken sowie zur Datenanalyse. Im Gegensatz zur Texterstellung und -bearbeitung sind solche Aufgaben in den bei der Befragung überwiegend vertretenden Branchen wenig verbreitet. Deshalb liegen bei diesen Anwendungszwecken die Anteile jener, bei denen die Systeme bereits zur Betriebspraxis gehören, zwischen zwölf und 19 Prozent. Bislang am wenigsten verbreitet und betrieblich im Einsatz sind mit Anteilen unter zehn Prozent generative KI-Systeme zur Erstellung und Bearbeitung von Videos sowie zur Ideenentwicklung und zum Ideenmanagement.

Abbildung 5: Betriebliche Nutzung von generativer KI

Frage: Für welche betrieblichen Zwecke werden in Ihrem Unternehmen bzw. Ihrer Verwaltung generative KI-Systeme wie ChatGPT genutzt oder ist ihre Nutzung geplant?



Dort, wo generative KI bislang noch keine Rolle spielt, bestehen jedoch vielfach Nutzungsabsichten für die Zukunft. Insbesondere bei der Übersetzung und bei der Datenanalyse werden von unseren Befragten noch große Nutzungspotenziale gesehen. Bei diesen Anwendungszwecken gibt jeweils etwa ein Fünftel an, dass der Einsatz der betreffenden Systeme für die Zukunft geplant

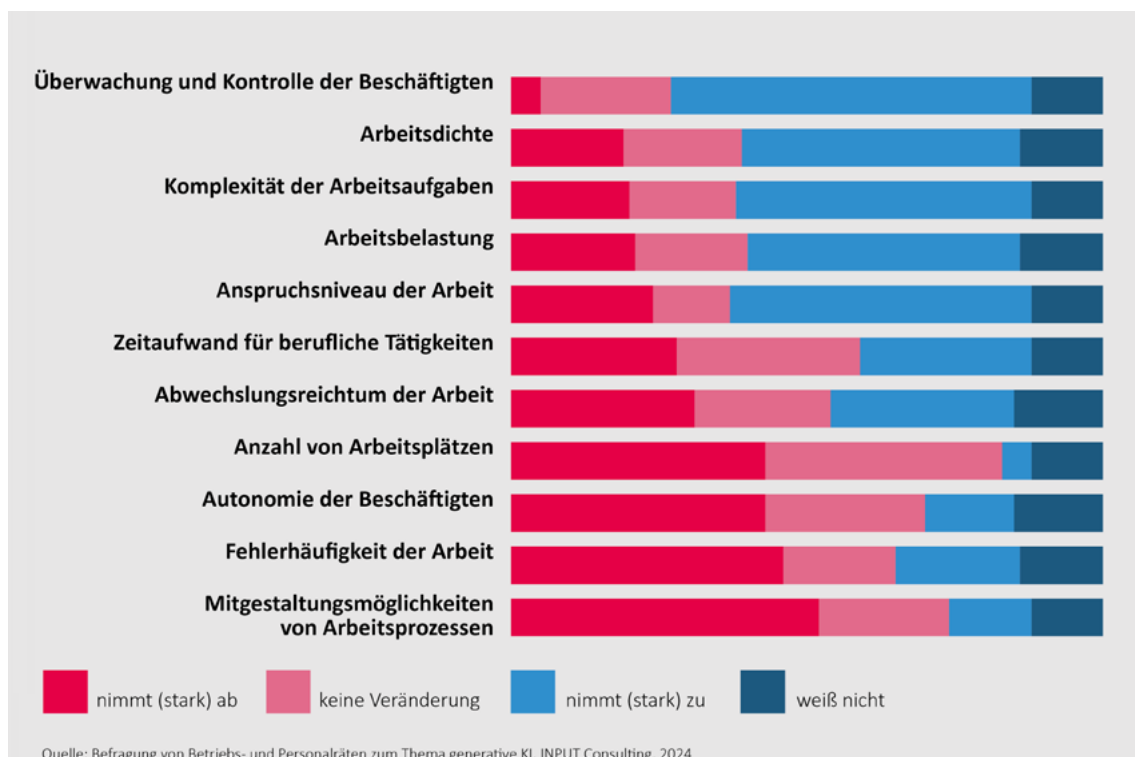
sei. Auffällig ist jedoch auch, dass die Anteile jener, die davon ausgehen, dass die entsprechenden generativen KI-Systeme auch in Zukunft keine Rolle spielen werden, mit Quoten zwischen 25 und 45 Prozent relativ hoch sind. Zudem deuten die hohen Anteile jener, welche die „weiß nicht“-Kategorie gewählt haben, darauf hin, dass generative KI noch nicht breitflächig in den Betriebs- und Personalratsgremien angekommen zu sein scheint, sondern sich bislang als Thema interessierter Spezialist*innen darstellt.

Welche Risikoeinschätzung zu derartigen Systemen ist bei Betriebs- und Personalräten vorhanden?

Hinsichtlich der erwarteten Auswirkungen des Einsatzes generativer KI-Systeme zeigt sich in den Antworten unserer Befragten eine gewisse Polarisierung in Richtung der negativen Beschäftigungsfolgen. So wird auf der einen Seite davon ausgegangen, dass generative KI die Arbeitssituation der Beschäftigten verschlechtern wird. Dies zeigt sich in der Befürchtung, dass mit generativen KI-Systemen eine stärkere Überwachung und Kontrolle der Beschäftigten verbunden ist. Davon gehen 61 Prozent der Befragten aus (siehe Abbildung 6). Zudem werden von knapp der Hälfte steigende Arbeitsbelastungen und eine erhöhte Arbeitsdichte infolge des Einsatzes generativer KI-Systeme erwartet. Ebenso dürfte generative KI nach Ansicht der Betriebs- und Personalräte zu einem Anstieg der Komplexität der Arbeitsaufgaben und damit einhergehend des Anspruchsniveaus der Arbeit führen. Auch hier ist etwa die Hälfte der Befragten der Meinung, dass sich die Arbeit in diese Richtung verändern werde. In der Frage, ob der Einsatz generativer KI auch Arbeitsplätze kosten wird, sind sich die Befragten uneinig. Knapp die Hälfte erwartet einen Rückgang an Jobs infolge der KI-Nutzung, etwa genauso viele gehen von keinen wesentlichen Änderungen der Anzahl an Arbeitsplätzen bzw. gar einer leichten Zunahme aus.

Abbildung 6: Beeinflussung der Beschäftigungssituation durch generative KI

Frage: In welche Richtung beeinflusst generative KI Ihrer Ansicht nach die Beschäftigungssituation der Mitarbeitenden in den nachfolgend genannten Aspekten?





Welche Folgen hat generative KI für die Arbeit?

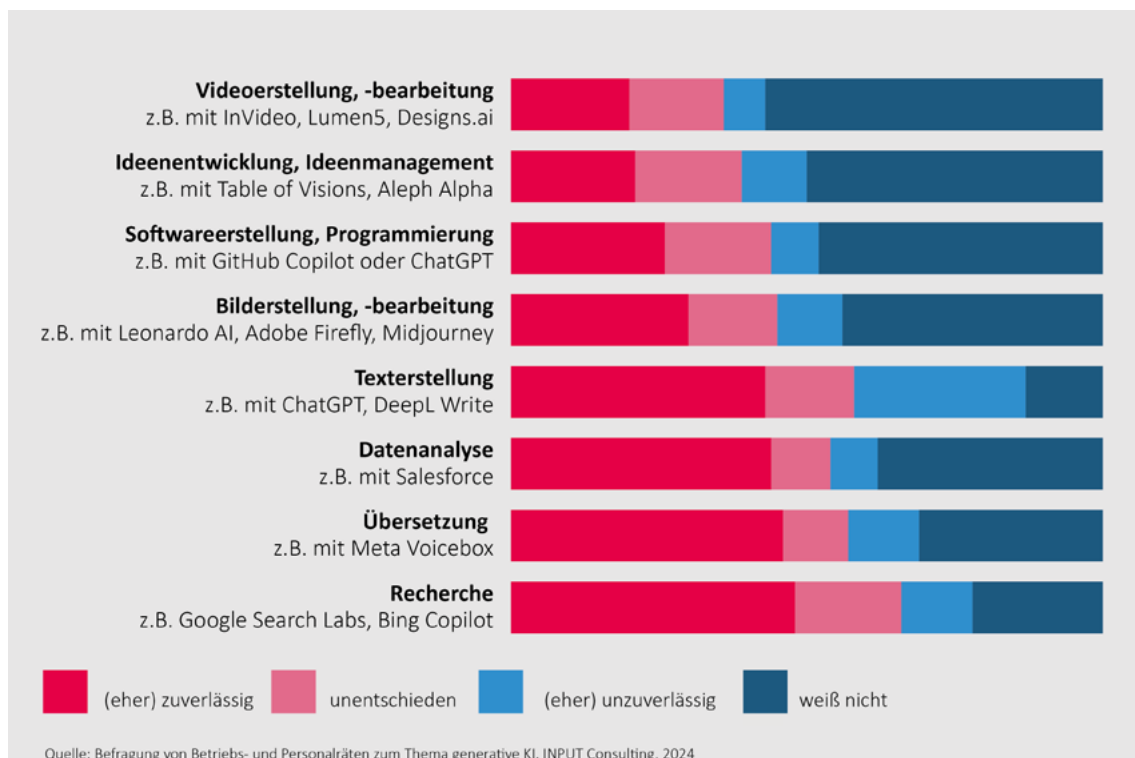
Aus unseren Ergebnissen wird jedoch deutlich, dass sich die Möglichkeiten zur Mitgestaltung von Arbeitsprozessen und auch die Autonomie in der Arbeit verringern dürften. Eine solche Sichtweise bringen jeweils knapp die Hälfte der Befragten zum Ausdruck. Auf der anderen Seite werden jedoch auch eine Verbesserung der Arbeitsbedingungen und der Arbeits- und Ergebnisqualität erwartet. So wird mehrheitlich davon ausgegangen, dass sich die Fehlerhäufigkeit bei der Arbeit reduziert. 47 Prozent der Befragten äußern eine entsprechende Erwartung, weitere 19 Prozent prognostizieren keine Veränderung. Eine gewisse Unentschiedenheit zeigt sich bei der Frage nach dem Zeitaufwand für berufliche Tätigkeiten. In diesem Aspekt halten sich die Anteile jener, die erwarten, dass dieser abnimmt (28 %), mit jenen die Waage, die von einer Zunahme ausgehen (29 %). Insgesamt deuten unsere Befragungsergebnisse auf ein relativ hohes Maß an Skepsis von Betriebs- und Personalräten hinsichtlich der Frage hin, ob der betriebliche Einsatz generativer KI-Systeme zum Vorteil der Beschäftigten erfolgt und er positive Auswirkungen auf die Beschäftigungssituation verspricht.

Welchen Einsatzgebieten wird welches Maß an Vertrauen entgegengebracht?

Für die Bewertung generativer KI spielt auch die Zuverlässigkeit der Anwendungen als ein wichtiges Vertrauenskriterium eine große Rolle. Unsere Befragungsergebnisse zeigen, dass generative KI-Systeme, die zu Recherchezwecken, zur Übersetzung, zur Datenanalyse und zur Texterstellung genutzt werden, im Urteil der Befragten überwiegend als zuverlässig gelten. Bei diesen Anwendungszwecken liegen die Anteile jener, die solchen Systemen Zuverlässigkeit attestieren, zwischen 40 und 50 Prozent. Bei generativer KI für die Nutzungszwecke Bilderstellung, Softwareerstellung und Programmierung, Ideenentwicklung sowie der Videoerstellung und -bearbeitung gibt es dagegen noch Vorbehalte.

Abbildung 7: Zuverlässigkeit des betrieblichen Einsatzes von generativer KI

Frage: Für wie zuverlässig halten Sie momentan den betrieblichen Einsatz generativer KI bei den nachfolgend genannten Nutzungszwecken?

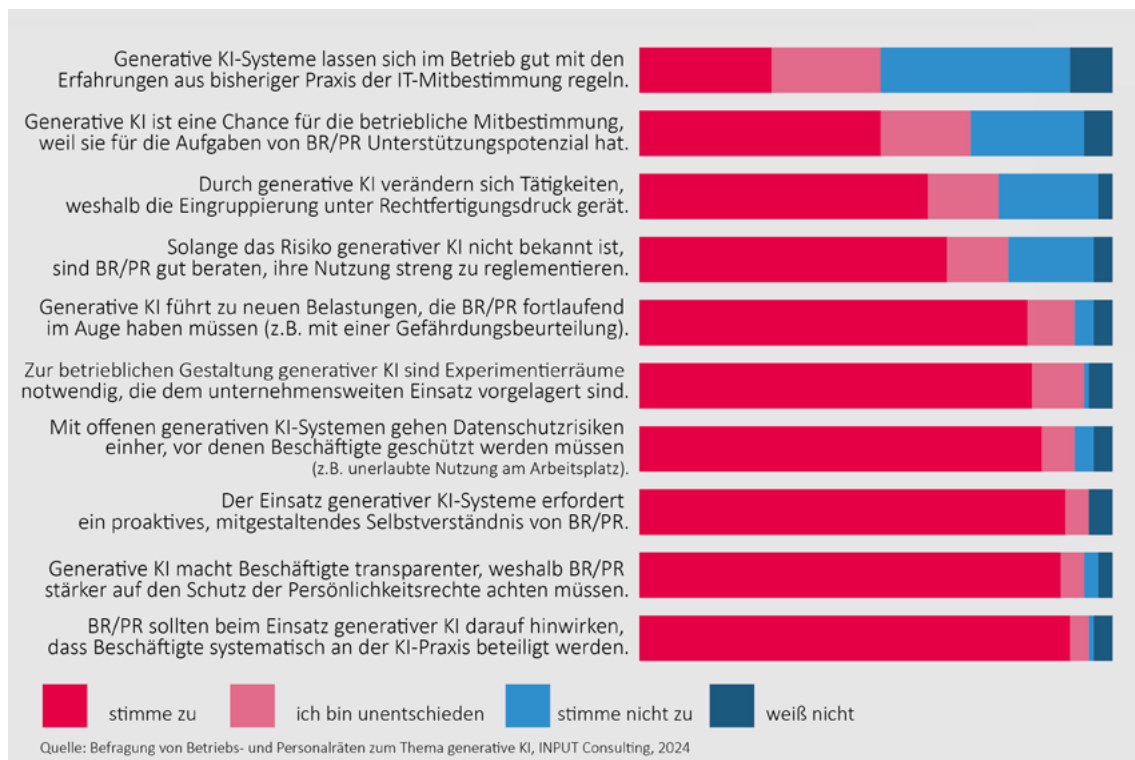


In den Antworten drückt sich jedoch auch eine gewisse Unerfahrenheit im Umgang mit generativer KI aus. Generative KI wird insbesondere dann als zuverlässig beurteilt, wenn eigene Nutzungserfahrungen vorzuliegen scheinen. Dafür sprechen auch die teils über 50 Prozent liegenden Anteile der „weiß nicht“-Antworten. Am belastbarsten scheinen die Erfahrungen bislang mit Systemen zur Texterstellung (z.B. mit ChatGPT) zu sein, weil die Befragten dort auch in höherem Maße von negativen Erfahrungen berichten, was bei den anderen Nutzungszwecken (aufgrund mangelnder Erfahrung und Praxis?) hingegen kaum der Fall ist.

Welche besonderen Herausforderungen stellt generative KI an Mitbestimmungsakteure?

Abbildung 8: Herausforderungen für die betriebliche Mitbestimmung bei Nutzung generativer KI

Frage: Welche besonderen Herausforderungen für die betriebliche Mitbestimmung sind mit generativen KI-Systemen verbunden?



Damit generative KI im Interesse der arbeitenden Menschen eingesetzt werden kann, bedarf es einer proaktiven, mitgestaltenden Mitbestimmungspraxis im Sinne einer prozessualen partnerschaftlichen Konfliktkultur.⁵⁴ Diese ist jedoch mit einer Reihe an Herausforderungen verbunden, weil es darum geht, die Sozialpartnerschaft neu auszurichten und die Mitbestimmung am Prozess zu orientieren, bevor über geeignete Regulierungsformen nachgedacht werden kann. Für eine solche prozessorientierte Mitbestimmungspraxis sind auch Kompetenzen in den Betriebs- und Personalratsgremien notwendig, die zunächst aufgebaut werden müssen.⁵⁵

54 Detlef Gerst, »Mitbestimmung in digitalen und agilen Betrieben – das Modell einer prozessualen partnerschaftlichen Konfliktkultur«, in: Verena Bader/Stephan Kaiser (Hg.), Arbeit in der Data Society, Wiesbaden 2020, S. 35-56.

55 Markus Hoppe/Ines Roth/Giovanni Suriano/Daniel Klueß, Künstliche Intelligenz menschenzentriert gestalten. Ein Praxisbericht aus dem Kundenservice, humain work lab KI-Lernraum 2022, <https://www.humain-worklab.de/wp-content/uploads/2022/04/KI-menschenzentriert-gestalten.-Ein-Praxisbericht-aus-dem-Kundenservice.pdf>

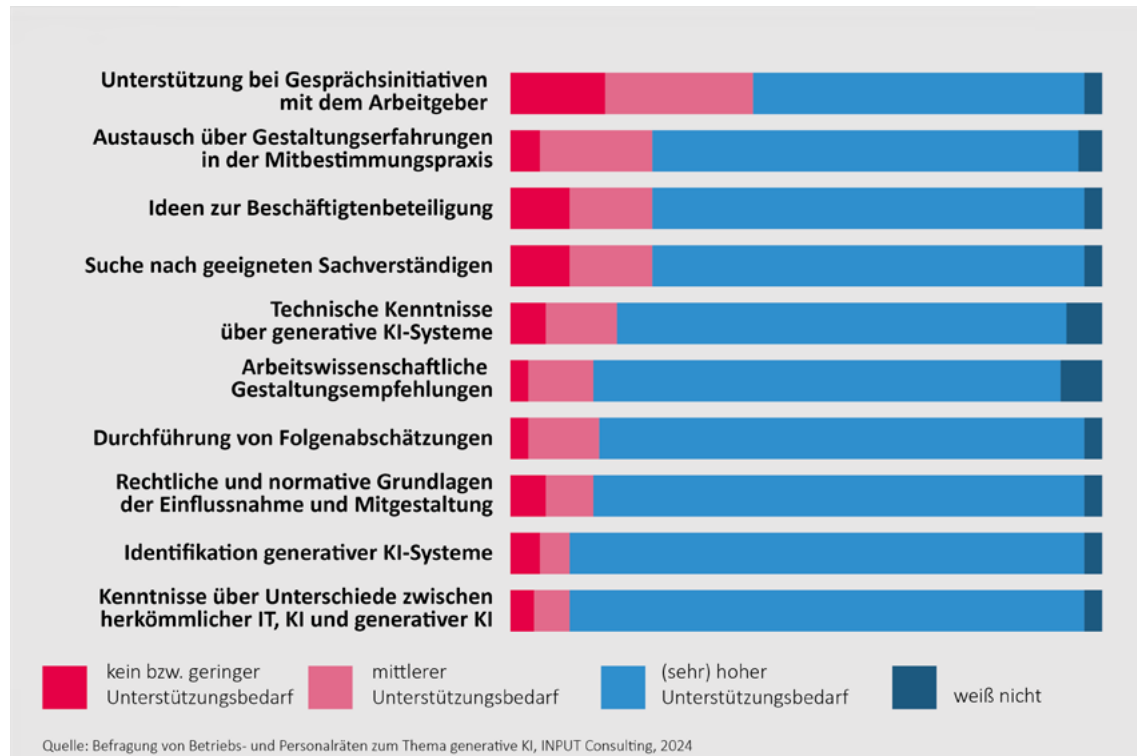


Vor diesem Hintergrund haben wir in unserer Befragung auch danach gefragt, welche Herausforderungen generative KI an Mitbestimmungsakteure stellt. Im Ergebnis zeigt sich, dass Betriebs- und Personalräte systematisch an der betrieblichen KI-Praxis beteiligt werden sollten und auch Beschäftigten eine entscheidende Rolle bei der Einführung der Systeme und der Weiterentwicklung der KI-Praxis zukommen sollte. Als eine Voraussetzung proaktiver, prozessorientierter Mitbestimmung erscheint aber, wie unsere Befragungsdaten auch zeigen, zunächst eine entsprechende strategische Positionierung der Gremien dem Thema KI gegenüber. Etwa 90 Prozent unserer Befragten teilen diese Einschätzung. Dass Beschäftigtenbeteiligung ein Kernelement der betrieblichen KI-Praxis sein soll, bestätigen mehr als 90 Prozent unserer Befragten. Als eine Möglichkeit, Beschäftigte zu beteiligen, gelten Experimentierräume, in denen vor dem flächendeckenden Einsatz KI-Systeme zunächst in einem sozialpartnerschaftlich geschützten Raum ausprobiert und hinsichtlich ihrer Wirkungen analysiert werden können. Mehr als 80 Prozent unserer Befragten sehen in Experimentierräumen ein geeignetes KI-Gestaltungsprinzip, das auch in der betrieblichen Mitbestimmung verankert sein sollte. Ebenso viele machen sich für den Schutz der Beschäftigten vor möglichen negativen Folgewirkungen des KI-Einsatzes (z.B. „gläserne Belegschaften“) stark. Als ein Thema der Mitbestimmung, das im Kontext des Einsatzes (generativer) KI hoher Aufmerksamkeit bedürfe, zeigt sich das Potenzial der KI-Systeme zur Veränderung von Tätigkeiten und Berufsbildern, wodurch die betriebliche Eingruppierungspraxis unter Druck geraten könne. Diese Perspektive wird von etwa zwei Dritteln unserer Befragten geteilt. Zudem stellt sich die Frage der Regulierung generativer KI-Systeme, in der sich aus dem Antwortverhalten der Befragten Handlungsbedarf ableiten lässt. Lediglich 28 Prozent sind der Meinung, dass sich generative KI-Systeme gut mit den Erfahrungen aus der bisherigen Praxis der IT-Mitbestimmung regeln lassen. Mit Blick auf die Potenziale, die generative KI auch für die Arbeit der Interessenvertretung entfalten kann, haben wir auch um eine Einschätzung darüber gebeten, inwieweit generative KI auch eine Chance für die betriebliche Mitbestimmung aufgrund ihres Unterstützungspotenzials für die Gremienarbeit sein kann. In dieser Frage herrscht jedoch Uneinigkeit vor, etwa die Hälfte geht von einem solchen Unterstützungspotenzial aus, knapp 20 Prozent sind unentschieden und etwa ein Viertel zeigt sich skeptisch. Insgesamt deuten die Ergebnisse aber darauf hin, dass es bei der betrieblichen Gestaltung und Mitbestimmung generativer KI auch für Betriebs- und Personalräte notwendig ist, die Veränderungen proaktiv zu bearbeiten, nah dran am Geschehen zu sein, sich für eine umfassende Beschäftigtenbeteiligung stark zu machen und die Risiken des KI-Einsatzes durch eine prozessorientierte Mitbestimmungspraxis im Blick zu haben und darauf reagieren zu können.

Welche Hilfe brauchen Mitbestimmungsträger*innen?

Abbildung 9: Unterstützungsbedarfe für Betriebs- und Personalräte bei der Nutzung von generativer KI

Frage: Wo sehen Sie Unterstützungsbedarfe für Betriebs- und Personalräte bei der Einführung und Nutzung von generativen KI-Systemen in den Unternehmen und Verwaltungen?



Generative KI ist für viele Betriebs- und Personalräte noch Neuland. Deshalb brauchen sie Unterstützung. Auffällig ist, dass die Unterstützungsbedarfe in nahezu allen der abgefragten Aspekte hoch sind. Zumeist bestehen Unklarheiten hinsichtlich der technischen, KI-spezifischen Aspekte generativer KI-Systeme und ihrer Identifikation. Kaum weniger wichtig erscheinen die rechtlichen und normativen Grundlagen der Einflussnahme und Mitgestaltung, die notwendig sind, um den betrieblichen Einsatz generativer KI als Gremium aktiv mitgestalten zu können. Ebenfalls hoch ist der Unterstützungsbedarf bei der Durchführung einer Folgenabschätzung, mit der ein Monitoring der Auswirkungen der KI-Nutzung für Beschäftigte gewährleistet werden kann. Bei den bislang genannten Aspekten geben jeweils mehr als 80 Prozent unserer Befragten an, einen sehr hohen Unterstützungsbedarf zu haben. Etwas weniger hoch erscheinen die Unterstützungsbedarfe im Hinblick auf arbeitswissenschaftliche Gestaltungsempfehlungen (etwa in Bezug auf eine menschenzentrierte KI-Anwendungspraxis im Betrieb), technische Kenntnisse über generative KI-Systeme, Ideen zur Beschäftigtenbeteiligung sowie den Austausch über Gestaltungserfahrungen in der Mitbestimmungspraxis. Bei diesen Aspekten geben jeweils etwa drei Viertel der Befragten an, einen hohen Unterstützungsbedarf zu haben. Am wenigsten Unterstützung benötigen die von uns befragten Betriebs- und Personalräte bei Gesprächsinitiativen mit dem Arbeitgeber in Sachen generativer KI. Hier artikuliert nur etwas mehr als die Hälfte einen hohen Unterstützungsbedarf, was zum einen daran liegen dürfte, dass dieser Aspekt auch in anderen Regulierungszusammenhängen bedeutsam ist und es deshalb Praxiserfahrungen gibt. Zum anderen können sich darin auch betriebskulturell begründete Vorbehalte gegen die Integration externer Unterstützer*innen in betriebliche Verhandlungsprozesse widerspiegeln.



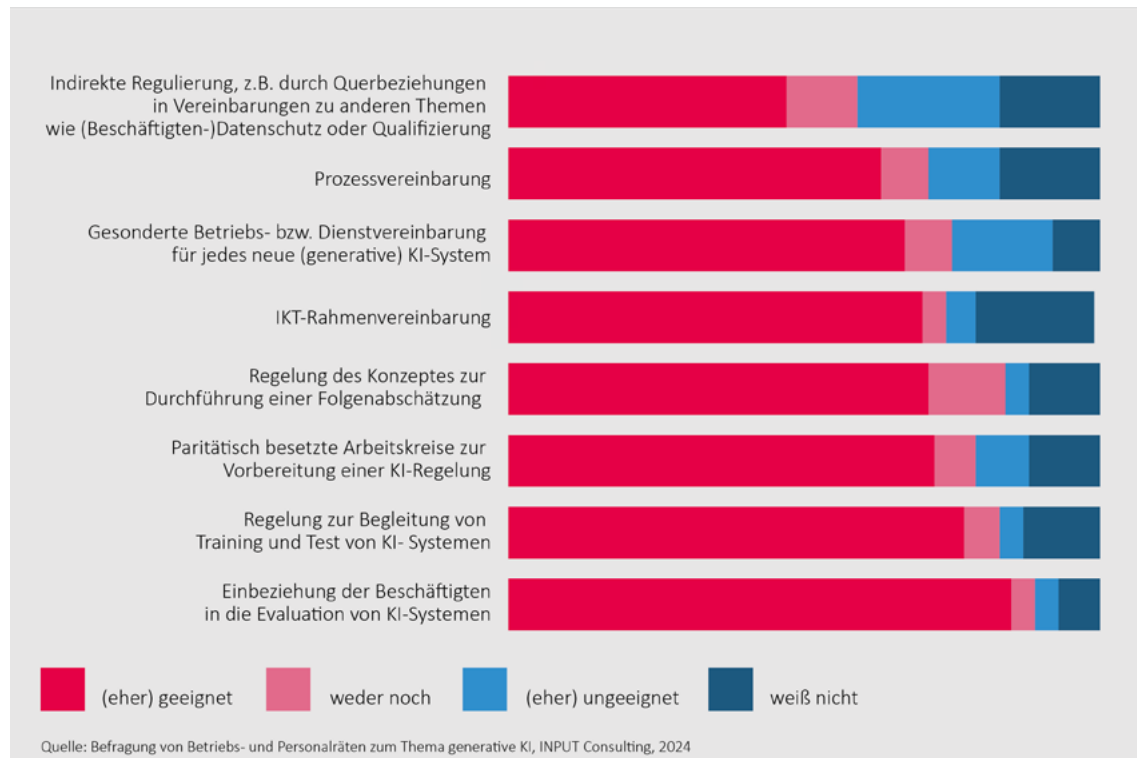
Als problematisch erweist sich zudem die Suche nach geeigneten Sachverständigen in KI-Fragen bei der Regulierung von (generativer) KI im Betrieb; 74 Prozent der Befragten äußern hier einen hohen Unterstützungsbedarf. Die Notwendigkeit zur betrieblichen Regulierung von KI resultiert auch aus der Komplexität von KI als Gegenstand der Betriebspolitik. Betriebs- und Personalrät*innen stoßen hier mitunter an ihre fachlichen Grenzen, weshalb der Gesetzgeber die Hinzuziehung externen Sachverständs ermöglicht, soweit dies zur Erfüllung der Aufgaben der Interessenvertretung notwendig ist (§ 80 Abs. 3 BetrVG).⁵⁶ Dieses Recht erfährt nun unter expliziter Bezugnahme auf Künstliche Intelligenz eine Stärkung, indem im Betriebsrätemodernisierungsgesetz klargestellt wird, dass nach § 80 Abs. 3 BetrVG die Hinzuziehung eines Sachverständigen durch den Betriebsrat dann als erforderlich gilt, wenn der Betriebsrat die Einführung oder Anwendung von Künstlicher Intelligenz bei der Durchführung seiner Aufgaben beurteilen muss. Mit Blick auf die im Betriebsrätemodernisierungsgesetz enthaltene Regelung, nach der Künstliche Intelligenz als Rechtfertigungsgrund für die Hinzuziehung externen Sachverständs anerkannt ist, haben wir in unserer Befragung 2024 deren praktische Relevanz überprüft. Dabei zeigt sich, dass insgesamt 21 Prozent aller Befragten angeben, dass bereits externer Sachverstand in KI-Fragen hinzugezogen worden ist. Zum überwiegenden Teil erfolgte dies aber unabhängig von der Regelung im Betriebsrätemodernisierungsgesetz; lediglich fünf Prozent aller Befragten geben an, sich auf diese gesetzliche Regelung berufen zu haben. Wenn externer Sachverstand hinzugezogen wurde, dann ging es zumeist um Hilfestellung bei der Formulierung konkreter Regulierungsgrundlagen oder um allgemeines Fachwissen hinsichtlich der Gestaltungsmöglichkeiten von KI-Systemen. Wenn beim KI-Einsatz kein externer Sachverstand hinzugezogen wurde, dann lag das am häufigsten daran, dass die Arbeitsschwerpunkte in den Gremien der Interessenvertretung nicht auf der KI-Regulierung lagen, die intern verfügbaren Kompetenzen für ausreichend gehalten wurden oder KI bislang noch keine Rolle im Betrieb gespielt hat. Zu einem geringen Teil sei auch die Regelung des Betriebsrätemodernisierungsgesetzes bislang nicht bekannt.

56 Deutscher Bundestag, »Gesetz zur Förderung der Betriebsratswahlen und der Betriebsratsarbeit in einer digitalen Arbeitswelt (Betriebsrätemodernisierungsgesetz)«, in: Bundesgesetzblatt 2021, S. 1762-1765, https://www.bgbl.de/xaver/bgbl/start.xav?startbk=Bundesanzeiger_BGBI&jumpTo=bgbl121s1762.pdf#__bgbl__%2F%2F*%5B%40attr_id%3D%27bgbl121s1762.pdf%27%5D__1664459476077

Welche betrieblichen Regulierungsvorgaben werden für erforderlich gehalten und welche existieren bereits?

Abbildung 10: Eignung von Instrumenten zur Regulierung von generativer KI

Frage: Welche Instrumente sind aus Ihrer Sicht für die Regulierung von generativen KI-Systemen in Unternehmen und Verwaltungen geeignet?



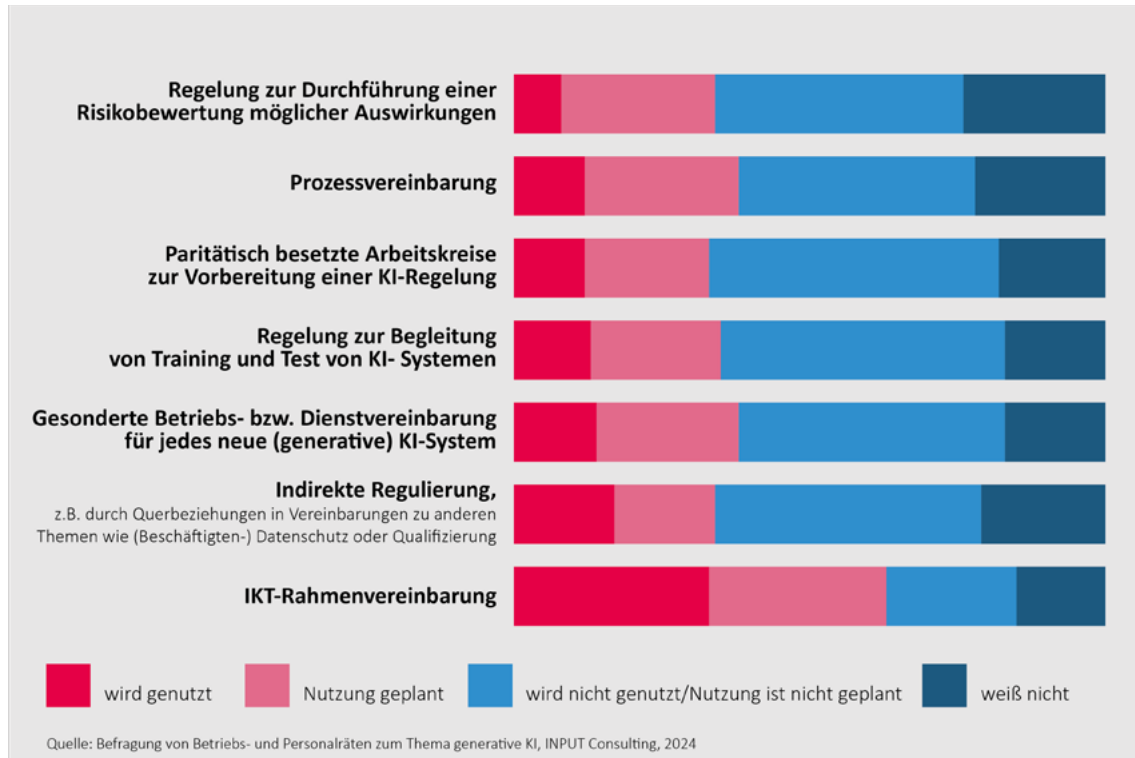
In unserer Befragung haben wir auch danach gefragt, welche Instrumente aus der Sicht von Betriebs- und Personalräten zur Regulierung generativer KI geeignet erscheinen. Dabei zeigt sich, dass dies insbesondere Instrumente sind, die auf eine Verbesserung der KI im Vorfeld und im laufenden Prozess abzielen und dabei die Nutzer*innen systematisch einbeziehen. So halten 86 Prozent der Befragten die „Einbeziehung von Beschäftigten in die Evaluation von KI-Systemen“ für ein geeignetes Regulierungsinstrument. Kaum weniger tauglich erscheint eine „Regelung zur Begleitung von Training und Test von KI-Systemen“ (76 %). Mit Anteilen zwischen 70 und 75 Prozent attestierter Eignung folgen „Paritätisch besetzte Arbeitskreise zur Vorbereitung einer KI-Regelung“, eine „Regelung zur Durchführung einer Folgenabschätzung“ und die „IKT-Rahmenvereinbarung“. In der Gunst der Befragten etwas abfallend sind die „gesonderte Vereinbarung jedes neuen (generativen) KI-Systems“ sowie die „Prozessvereinbarung“. Diese Instrumente gelten bei 67 bzw. 64 Prozent als geeignet. Am wenigsten zur Regulierung generativer KI tauglich ist im Urteil der Befragten die „indirekte Regulierung“, bei der generative KI durch Querverweis in Regelungen zu anderen, verwandten Themen „mitgeregelt“ wird. Dies halten nur 47 Prozent der Befragten für geeignet bei der Regulierung generativer KI.



Welche Folgen hat generative KI für die Arbeit?

Abbildung 11: Nutzung von Instrumenten zur Regulierung von generativer KI

Frage: Welche Instrumente zur Regulierung von generativen KI-Systemen werden in Ihrem Unternehmen bzw. Ihrer Verwaltung eingesetzt? Hinweis: Nur Befragte, bei denen generative KI zum Einsatz kommt oder der Einsatz geplant ist



Doch welche Instrumente werden bislang zur Regulierung generativer KI tatsächlich genutzt? Wenngleich die IKT-Rahmenvereinbarung wie gezeigt nur im Mittelfeld der als geeignet eingeschätzten Instrumente rangiert, ist diese Form bislang bei der Regulierung generativer KI am weitesten verbreitet. Etwa ein Drittel der Befragten, bei denen generative KI eine Rolle spielt, integrieren diese in den Geltungsbereich einer IKT-Rahmenvereinbarung, weitere 30 Prozent planen, dies in Zukunft zu tun. Auch die indirekte Regulierung, in der Liste der geeigneten Instrumente auf dem letzten Platz liegend, spielt in der Praxis eine weitaus größere Rolle. Aktuell nutzen 17 Prozent diese Regulierungsform. Die übrigen Regulierungsinstrumente werden zu vergleichbaren Anteilen von acht bis 14 Prozent genutzt. Somit ist insgesamt ein gewisser Pragmatismus im Umgang mit der Regulierungsfrage generativer KI erkennbar, der sich dadurch ausdrückt, dass – wissend, dass es besser geeignete Alternativen geben könnte – vor allem Regulierungsinstrumente genutzt werden, die es entweder bereits im Unternehmen bzw. der Verwaltung gibt oder die sich mit überschaubarem Aufwand schaffen lassen.

WIE LÄSST SICH
GENERATIVE KI IM
BETRIEB GESTALTEN?





WIE LÄSST SICH GENERATIVE KI IM BETRIEB GESTALTEN?

Praxistipps für Betriebs- und Personalräte

Die betriebliche Anwendung von ChatGPT und ähnlicher Modelle braucht betriebliche Regeln. Betriebs- und Personalräte sollten ihre Möglichkeiten der Einflussnahme nutzen, um interne Bedingungen zu formulieren und Leitlinien für die Nutzung entsprechender Tools zu schaffen. Dazu bieten sich folgende Tipps für das Wollen und Handeln an. Aber Vorsicht: Nicht jede Empfehlung passt zu jeder Anwendung, jeder Betriebskultur, jedem Einsatzzweck. Aktive Einflussnahme sollte mit Priorisierung beginnen:

Erstinformation/Meinungsbildung

- Lassen Sie sich erläutern, warum der Einsatz generativer KI als Werkzeug für die vorgesehenen Zwecke nützlich und erforderlich ist.
- Hinterfragen Sie, woher die ursprünglichen Trainingsdaten stammen und welche Einsatzreferenzen zum KI-Modell existieren. Damit lässt sich ein Eindruck darüber gewinnen, welche Weltsicht und welches Wertemodell den grundlegenden Funktionsmechanismen des Systems zugrunde lagen und welche Erfahrungen damit gemacht wurden.
- Fragen Sie danach, ob personenbezogene Daten, die von Dritten stammen, von diesen rechtmäßig erhoben worden sind und eine Rechtsgrundlage dafür existiert, diese Daten an das KI-System weiterzugeben. Akzeptieren Sie nicht die Behauptung, dass alle im Netz verfügbaren Daten auch automatisch legal verarbeitbar sind.
- Lassen Sie sich Informationen darüber geben, wie die einzelnen Anwendungen generativer KI in die Risikoklassifikation der europäischen KI-Verordnung eingeordnet und wie die Anforderungen der Verordnung erfüllt werden.
- Legen Sie Wert darauf, über die Auswirkungen der Anwendungen auf die Persönlichkeitsrechte nach der europäischen Grundrechtecharta informiert zu werden. Drängen Sie auf eine Grundrechtsfolgeabschätzung, zumindest für heikle Anwendungen, z. B. im Hochrisikobereich.
- Stellen Sie fest, ob der Einsatz der KI-Anwendung Auswirkungen hat auf die Arbeitssteuerung im Betrieb. Falls ja: Reklamieren Sie eine Folgeabschätzung für würdige, sichere und gesunde Arbeitsbedingungen, die mit Art. 31 der europäischen Grundrechtecharta und mit deutschem Arbeitsschutzrecht begründet werden kann.
- Fragen Sie danach, inwieweit der Einsatz generativer KI herkömmliche Haftungsprinzipien im Betrieb beeinflusst und wie der Betrieb mit eventuell entstehenden urheberrechtlichen Konflikten umgehen will.

Dialog

- Interessieren Sie sich für die Meinung der Beschäftigten zum Einsatz generativer KI. Thematisieren Sie in einer Betriebsversammlung Chancen und Risiken. Beteiligen Sie Beschäftigte daran, eine Position zur Regulierung dieser KI-Systeme zur Festlegung von „roten Linien“ zu formulieren.
- Entwickeln Sie am besten gemeinsam mit dem Arbeitgeber Regeln, die beschreiben, für welchen Einsatzzweck generative KI in ihrem Betrieb eingesetzt werden darf und für welchen nicht. Beschreiben Sie ihre „rote Linien“ und vermitteln Sie diese dem Arbeitgeber.
- Entwickeln Sie am besten gemeinsam mit dem Arbeitgeber ein Standardverfahren zur Befassung der Mitbestimmungsorgane mit Anwendungen, die auf generativer KI beruhen. Ein gemeinsam formulierter Qualitätsanspruch erleichtert das Handeln. Darin sollten Qualitätsfaktoren wie Diskriminierungsfreiheit, Rechtskonformität und Nachhaltigkeit ebenso wie Fragen der Ergonomie, der Sozialverträglichkeit, der Robustheit der Systeme sowie die Unterstützung von Ansprüchen an gute Arbeit und Beschäftigung zum Maßstab erhoben werden. Einkauf, Entwicklung und Einsatz generativer KI-Systeme sollte auf derartige Qualitätsfaktoren ausgerichtet werden.

Besser abgeschottet als offen

- Priorisieren Sie den Einsatz eines Systems, das das Unternehmen unter eigener Kontrolle hat, statt ein offenes und frei zugängliches System zu verwenden. Derartige geschlossene Systeme sind zwar vortrainiert, ihre wichtigen Daten bleiben aber in der Firma. Damit werden die Nutzungsgewohnheiten nicht zum weiteren Training des Systems an den Anbieter der generativen KI weitergegeben. Das schützt die Daten von Beschäftigten und die vertraulichen Informationen der Firma.
- Legen Sie Wert darauf, dass berufliche Accounts für die Nutzung generativer KI zur Verfügung gestellt werden. Damit wird vermieden, dass Beschäftigte eigenverantwortlich unter Verwendung privater Daten ein Konto erstellen, das beruflich genutzt wird. Ihr persönliches Nutzungsverhalten bleibt damit für die Anbieter der Software schlechter auswertbar und zuzuordnen.

Transparenz

- Legen Sie Wert darauf, dass maschinell erzeugte Texte und Programmcodes, die betrieblich verwendet werden sollen, als „computergeneriert“ gekennzeichnet werden, sofern diese zuvor nicht von Verantwortlichen überprüft, verändert und nachvollziehbar freigegeben wurden. Die Verantwortlichen für den maschinellen Output übernehmen nach der Freigabe die Verantwortung für den Text bzw. Programmcode.
- Wirken Sie darauf hin, dass künstlich erzeugte Bilder, die Personen abbilden oder Beschäftigte unerschwellig beeinflussen könnten, auf jeden Fall als „computergeneriert“ gekennzeichnet werden.
- Setzen Sie sich dafür ein, dass für die Beschäftigten Transparenz darüber geschaffen wird, welche Anwendungen generativer KI im Betrieb zugelassen wurden. Drängen Sie darauf, dass auch die dafür geltenden Regelungen für die Beschäftigten verständlich beschrieben und leicht zugänglich sind. Beschäftigte sollten daran erinnert werden, welche Interventionsmechanismen sie bei der Verwendung ihrer personenbezogenen Daten haben und darüber aufgeklärt werden, welche Art von betrieblichen Anwendungen nach der europäischen KI-Verordnung dem Hochrisikobereich zugeordnet wurden.
- Weisen Sie die Beschäftigten darauf hin, dass nicht regelgerechte Verwendung von generativer KI die Persönlichkeitsrechte der Mitarbeiter gefährden kann und alle Beschäftigten sowohl Interventionsmöglichkeiten als auch die Möglichkeit haben, Hinweise über Regelverstöße an die Mitbestimmungsgremien oder die Hinweisgeberportale des Arbeitgebers weiterzugeben. Erinnern Sie daran, dass die Whistleblower nach dem Hinweisgeberschutzgesetz besonders geschützt sind.

Rote Linien

- Begründen Sie den Anspruch, grundsätzlich keine weiteren personenbezogenen Daten an das System zu übergeben und keine personenbezogenen Daten abzufragen. Werden, nicht regelgerecht, trotzdem Daten über einzelne Beschäftigten durch das System zur Verfügung gestellt, sollte diese Information den Betroffenen zur Kenntnis gegeben werden, damit diese ihre Widerspruchsrechte wahrnehmen können. Derartige maschinelle Ausgaben sollen nicht für arbeitsrechtliche Maßnahmen und Leistungs- oder Verhaltenskontrolle genutzt werden dürfen.
- Lassen Sie nicht zu, dass Mithilfe von generativer KI automatisierte Entscheidungen mit Rechtswirkung oder ähnlicher Tragweite für Beschäftigte getroffen werden. Erinnern Sie an den Art. 22 der Datenschutzgrundverordnung.
- Fordern Sie ein, dass durch die eingeführten Anwendungen gewerkschaftliches Engagement oder Handeln der Mitbestimmungsorgane nicht ausgewertet werden und keine Informationen über die handelnden Akteur*innen analysiert werden. Erinnern Sie an die Vorgaben des Art. 9 der Datenschutzgrundverordnung, wonach die Verarbeitung weiterer besonders geschützter Daten untersagt ist.

Reglements

- Verlangen Sie, falls der Einsatzzweck die Verwendung personenbezogener Daten zwingend erfordert, dass diese Daten nur in mitbestimmten und streng regulierten Anwendungsszenarien verwendet werden dürfen. Drängen Sie die Verantwortlichen darauf, ihnen für solche Fälle eine Rechtsgrundlage zu nennen. Zunächst ist aber die Möglichkeit zu prüfen, ob der Zweck nicht auch mit pseudonymisierten oder anonymisierten Daten erfüllt werden kann.
- Setzen sie sich dafür ein, dass die Chathistorie für die beruflichen Accounts nur den Nutzer*innen zugänglich ist und eine angemessen kurze Speicherfrist dafür sorgt, dass die Daten über das Abfrageverhalten der Beschäftigten nicht unnötig lange im System vorgehalten werden.



Werte

- Übersetzen Sie im Betrieb ohnehin vorhandene Ethikprinzipien in Anwendungsvorgaben für die generative KI. Verlangen Sie, dass dabei geeignete, schon existierenden Ansprüche an Umgangsformen, Gesetzestreue, Vielfalt, Fairness und Arbeitnehmerrechte in die betrieblichen Regeln zum Einsatz generativer KI übertragen werden.
- Legen Sie Wert darauf, dass seitens des Anbieters der generativen KI verhindert wird, dass Abfragen zu sexuellen Inhalten, Gewalt und Hetze ermöglicht werden.

Prozesse

- Wirken Sie darauf hin, dass ein betriebliches Meldesystem aufgebaut wird, über das Beschäftigte auf eine nicht regelgerechte Nutzung der Anwendungen hinweisen können. Hinweisgeber sollten unbeschadet eine verantwortliche Stelle darauf aufmerksam machen können, wenn mit dem System Falschinformationen, Hassnachrichten, Fake News, manipulative Inhalte, Schadsoftware oder verleumderische Informationen über einzelne Beschäftigte verbreitet werden und Persönlichkeitsrechte der Beschäftigten bedroht sind.
- Drängen Sie darauf, dass Modelltests durchgeführt werden, bevor Anwendungen für den innerbetrieblichen Einsatz freigegeben werden. Motivieren Sie Mitbestimmungsträger*innen, sich an den Modelltests aktiv zu beteiligen, um eigene Eindrücke sammeln zu können, und fragen Sie beteiligte Beschäftigte nach deren Erfahrungen.
- Begründen Sie den Anspruch, dass bei erkennbaren negativen Auswirkungen des Einsatzes generativer KI auf Arbeitsplätze, Qualifikationsanforderungen und die Eingruppierung von Beschäftigten eine Grundrechtsfolgeabschätzung durchgeführt wird.
- Unterstützen Sie Initiativen, die darauf zielen, die Bedingungen und Wirkungen hinsichtlich der schon eingesetzten generativen KI-Anwendungen in ihrem Betrieb laufend zu evaluieren.
- Unterstützen Sie Anstrengungen, Mechanismen zu entwickeln, die Informationen darüber liefern, ob die im Betrieb eingesetzten Texte, Bilder und Programmcodes computergeneriert wurden oder nicht. Falls geeignete elektronische Prüftools dafür verfügbar werden, stehen Sie deren Einsatz aufgeschlossen gegenüber.

Kein blindes Vertrauen

- Hinterfragen Sie, ob den Entscheidungsträger*innen über den Einsatz generativer KI in ihrem Betrieb klargemacht wurde, dass mangelnde Genauigkeit, systembedingte Verzerrungen oder Neigungen dazu, unsinnige und falsche Inhalte auszugeben, die Verwendung entsprechender Systeme für manche Anwendungsszenarien nicht zulässt.
- Begründen Sie die Pflicht für Nutzer*innen von generativer KI in ihrem Betrieb, den maschinellen Output auf Richtigkeit, Stimmigkeit, Integrität, Genauigkeit und Diskriminierungsfreiheit zu prüfen, bevor er in den innerbetrieblichen Einsatz gebracht wird.
- Arbeiten Sie darauf hin, dass den Entscheidungsträger*innen vermittelt wird, nicht blind auf Empfehlungen der KI zu vertrauen und sich nicht bedingungslos an KI-gestützte Vorschläge gebunden zu fühlen. Fragen Sie bei der Evaluation der Systeme danach, wie häufig die Entscheidungsträger*innen von den maschinellen Empfehlungen abweichen.

Kompetenzen

- Interessieren Sie sich dafür, ob sich das Management und die operativen Entscheidungsträger*innen über die Vorgaben der europäischen KI-Verordnung und deren Sanktionsklauseln im Klaren sind.
- Setzen Sie sich für Qualifizierungsmaßnahmen ein, die den Anwender*innen vermitteln, welche Stärken und welche Schwächen generative KI hat, wie Mitarbeiter*innen die Modelle nutzen können und welche betrieblichen Regulierungsvorgaben es dafür gibt.

Technisch-organisatorische Sicherung

- Hinterfragen Sie, ob differenzierte Regulierungsvorgaben für den Einsatz generativer KI in den verschiedenen Anwendungsfeldern existieren.
- Achten Sie darauf, dass Datenschutzbeauftragte an der Formulierung der betrieblichen Regelungen beteiligt werden, und prüfen, ob ausreichend technisch-organisatorische Maßnahmen den sicheren Betrieb gewährleisten.
- Unterstützen Sie Anstrengungen, dezidierte Berechtigungskonzepte zu erstellen und technisch-organisatorische Schutzmaßnahmen zu integrieren, die die Anwendungen vor Missbrauch und unerlaubten Eingriffen schützen.

CHECKLISTE

- ☐ Mit welchen Daten wurde die generative KI trainiert?
- ☐ Welche Firma betreibt das KI-System? Wo sind die Daten gespeichert, die die Beschäftigten in das System eingeben?
- ☐ Soll ein KI-System eingesetzt werden, dass unter der vollen Kontrolle des Unternehmens ist, oder soll ein freier Zugang zu einer allgemein zugänglichen generativen KI organisiert werden?
- ☐ Ist es zulässig, personenbezogene Daten in dem System zu verarbeiten? Auf welcher Rechtsgrundlage erfolgt dies?
- ☐ Falls die generative KI personenbezogene Daten der Beschäftigten nutzen darf, gibt es dafür eine Zustimmung der Mitbestimmungsorgane?
- ☐ Liegen aus anderen Betrieben Angaben und Referenzen über die Verzerrungs- und Diskriminierungsfreiheit des Systems vor?
- ☐ Welche Erfahrungen und Beschreibungen zum Genauigkeitsgrad des Systems gibt es?
- ☐ Wurde das System in der Firma getestet, und welche Neigung dazu, unsinnige und falsche Ergebnisse auszugeben, wurde dabei festgestellt?
- ☐ Wurden die Nutzer*innen und Vorgesetzten darüber informiert, welche Unzulänglichkeiten das System hat? Wurde ihnen vermittelt, dass sie sich nicht hundertprozentig auf die Ergebnisse verlassen können?
- ☐ Liegen Erfahrungen darüber vor, inwieweit die Beschäftigten den Systemansatz für aufgabenangemessen und gebrauchstauglich halten?
- ☐ Sind die Beschäftigten über Chancen und Risiken informiert worden und welche Regeln für die Systemnutzung halten sie für besonders bedeutsam?
- ☐ Sollen vorausschauend die geplanten Vorgaben der europäischen KI-Verordnung zur Transparenz der generativen KI eingehalten werden?
- ☐ Wird das System in einem Anwendungsfeld eingesetzt, das nach der KI-Verordnung als Hochrisikobereich definiert ist?
- ☐ Wurde eine Grundrechtsfolgeabschätzung zum Einsatz des Systems durchgeführt? Mit welchem Ergebnis?
- ☐ Hat der Systemeinsatz Auswirkungen auf die Arbeitssteuerung im Betrieb und wurden Konsequenzen für Arbeitsschutz und Persönlichkeitsrechte in die Folgeabschätzung einbezogen?
- ☐ Wie werden die personenbezogenen Daten der Beschäftigten im Betrieb geschützt? Dürfen derartige Daten durch das System ausgewertet werden? Wer hat auf die Auswertungen Zugriff?
- ☐ Gibt es eine klare „Gebrauchsanweisung“, die beschreibt, für welche Zwecke derartige Systeme im Betrieb eingesetzt werden dürfen und für welche nicht?
- ☐ Haben nur die Nutzer*innen Zugang zur Historie der von ihnen eingegebenen Prompts? Wie lange bleiben die Eingabedaten gespeichert?



Wie lässt sich generative KI im Betrieb gestalten?

- ☐ Existieren betriebliche Vorgaben, dass Texte, Bilder und Programmcodes, die mit generativer KI erzeugt werden, als „computergeneriert“ zu kennzeichnen sind?
- ☐ Werden künstlich erzeugte Bilder, die Personen aus der Firma nachempfunden sind, als computergeneriert gekennzeichnet?
- ☐ Sind die Regeln, die für den Einsatz generativer KI gelten, für die Beschäftigten zugänglich und verständlich geschrieben?
- ☐ Wurde den Nutzer*innen und den Vorgesetzten vermittelt, welche Interventionsmöglichkeiten Betriebsangehörige haben, um auf Regelverstöße bei der Verwendung der generativen KI hinzuweisen?
- ☐ Werden mithilfe generativer KI vollständig automatisierte Entscheidungen getroffen, die auf Beschäftigte wirken?
- ☐ Ist es im Betrieb zulässig, das Verhalten von Mitbestimmungsorganen durch generative KI auswerten zu lassen?
- ☐ Wurde überprüft, wie der Einsatz generativer KI zu den Ethikprinzipien der Firma passt und wurden aus dem Abgleich Schlussfolgerungen für die betrieblichen Vorgaben gezogen?
- ☐ Wurden betriebliche Modelltests durchgeführt, bevor das KI-System zum dauerhaften Einsatz kam? Wurden die Beschäftigten und ihre Interessenvertretungen daran beteiligt?
- ☐ Wurde analysiert, wie der Einsatz generativer KI auf die Anzahl der Arbeitsplätze, die Veränderung von Qualifikationsanforderungen und die Eingruppierung von Beschäftigten wirkt?
- ☐ Haben die Nutzer*innen des Systems die Vorgabe, den maschinellen Output auf Richtigkeit, Stimmigkeit, Integrität und Genauigkeit sowie auf Diskriminierungsfreiheit zu prüfen?
- ☐ Haben die Nutzer*innen Zugang zu Qualifizierungsmaßnahmen, die ihnen vermitteln, wie man zielführende Prompts schreibt? Wurde den Nutzer*innen vermittelt, welche Regeln für den Einsatz der generativen KI im Betrieb gelten?
- ☐ Wurde innerbetrieblich geprüft, inwieweit für die unterschiedlichen Einsatzgebiete generativer KI (zum Beispiel Softwareentwicklung oder Bildgenerierung) unterschiedliche Regulierungsvorgaben notwendig sind?

☐ _____

☐ _____

☐ _____

☐ _____

☐ _____

IN WELCHEN RECHTSRAHMEN
IST DER EINSATZ
GENERATIVER KI
EINGEBETTET?





IN WELCHEN RECHTSRAHMEN IST DER EINSATZ GENERATIVER KI EINGEBETTET?

Die rechtliche Basis der KI-Gestaltung

Kaum war die europäische KI-Verordnung in Kraft, war die Presse schon voll von „Wenns“ und „Abers“ zur Notwendigkeit, dem Einsatz Künstlicher Intelligenz einen Regulierungsrahmen zu geben. Vernachlässigt wird in solchen öffentlichen Debatten häufig, dass schon zahlreiche Gesetze existieren, die auch auf den Einsatz von KI-Systemen wirken. Das Bundesdatenschutzgesetz, die europäische Datenschutzgrundverordnung, das Hinweisgeberschutzgesetz, das Allgemeine Gleichbehandlungsgesetz und andere geben bereits einen Handlungsrahmen vor. Das Betriebsverfassungsgesetz schafft die Möglichkeit, auf den betrieblichen Einsatz von KI-Systemen Einfluss zu nehmen.

Die nachstehende Übersicht erhebt nicht den Anspruch, im Sinne eines Rechtskommentars in die Details dieser Rechtsnormen einzuführen. Sie will eine Übersicht über Rechtsgrundlagen geben und grob über den europäischen Rechtsrahmen für KI-Systeme informieren. Die Auswahl der zitierten Rechtsvorgaben und Gesetzentwürfe ist im weitesten Sinne auf die Möglichkeiten ausgerichtet, Qualitätsmaßstäbe für generative KI zu formulieren und entsprechend ihrem Anwendungsgebiet darauf Einfluss zu nehmen. Es bleibt eine Übersicht.

Auf die umfassende Darstellung auch existierender und prägender Normen, der Vorgaben von Aufsichtsbehörden, Branchenverbänden, Selbstverpflichtungen und branchenspezifischen Gesetzen musste zu Gunsten der Schwerpunktsetzung dieser Broschüre verzichtet werden. Leser*innen, die tiefer ins Thema einsteigen möchten, seien auf das „Praxishandbuch Künstliche Intelligenz“ von Lothar Schröder und Petra Höfers verwiesen, das im Bund Verlag erschienen ist. Darin wird das Ordnungsprinzip „KI Lagom“ aus einer umfassenden Analyse weiterer Regelwerke abgeleitet.

Bundesdatenschutzgesetz

Das deutsche Bundesdatenschutzgesetz dient dem Schutz personenbezogener Daten und bleibt auch nach dem Inkrafttreten der europäischen Datenschutz-Grundverordnung eine bedeutsame Vorgabe. Hiernach kann ein Tarifvertrag oder eine Betriebs- bzw. Dienstvereinbarung eine Rechtsgrundlage bilden, um personenbezogenen Daten von Beschäftigten zu verarbeiten. Betriebsräte nutzen diese gesetzliche Möglichkeit in Kombination mit ihren Mitbestimmungsrechten gern, um Schutzregeln für die Beschäftigten festzulegen, wenn deren Daten verarbeitet werden. Dies ist auch möglich, wenn Systeme eingesetzt werden, die mit Künstlicher Intelligenz arbeiten. Mittels Betriebs- und Dienstvereinbarungen können Spielregeln auch für den Einsatz generativer KI im Betrieb verabredet werden.

§ 26 BDSG
Datenverarbeitung für Zwecke des Beschäftigungsverhältnisses

„Personenbezogene Daten (dürfen) verarbeitet werden, wenn dies [...] zur Ausübung oder Erfüllung der sich aus einem Gesetz oder einem Tarifvertrag, einer Betriebs- oder Dienstvereinbarung (Kollektivvereinbarung) ergebenden Rechte und [...] erforderlich ist.“

Siehe auch Artikel 6 und 88 DSGVO

Datenschutzgrundverordnung

Obwohl die europäische Datenschutz-Grundverordnung (DSGVO) seit Jahren gilt, wird ihr Eingangssatz offenbar häufig nicht gelesen. Darin steht, dass diese Verordnung die Grundrechte und Grundfreiheiten der Menschen in der EU und deren Daten schützt. Die EU-Grundrechtscharta wird damit schon im laufenden Recht eingefordert. Die gültige Datenschutz-Grundverordnung macht aber u.a. auch Vorgaben zum Schutz besonderer Kategorien personenbezogener Daten. Sie definiert beispielsweise, was unter Profiling zu verstehen ist und macht Vorgaben zur Datensicherheit und zur Datenschutzfolgenabschätzung.

Artikel 1 Abs 2 DSGVO	„Diese Verordnung schützt die Grundrechte und Grundfreiheiten natürlicher Personen und insbesondere deren Recht auf Schutz personenbezogener Daten.“ (Verweis auf die Europäische Grundrechtscharta)
Artikel 9 DSGVO Besondere Kategorien personenbezogener Daten	„Die Verarbeitung personenbezogener Daten, aus denen die rassische und ethnische Herkunft, politische Meinungen, religiöse oder weltanschauliche Überzeugungen oder die Gewerkschaftszugehörigkeit hervorgehen, sowie die Verarbeitung von genetischen Daten, biometrischen Daten zur eindeutigen Identifizierung einer natürlichen Person, Gesundheitsdaten oder Daten zum Sexualleben oder der sexuellen Orientierung einer natürlichen Person ist untersagt.“
Artikel 22 und 4 Nr. 4 DSGVO Automatisierte Entscheidungen und Profiling	„Die betroffene Person hat das Recht, nicht einer ausschließlich auf einer automatisierten Verarbeitung – einschließlich Profiling – beruhenden Entscheidung unterworfen zu werden, die ihr gegenüber rechtliche Wirkung entfaltet oder sie in ähnlicher Weise erheblich beeinträchtigt.“ „[...] der Verantwortliche trifft angemessene Maßnahmen, um die Rechte und Freiheiten sowie die berechtigten Interessen der betroffenen Person zu wahren, wozu mindestens das Recht auf Erwirkung des Eingreifens einer Person seitens des Verantwortlichen, auf Darlegung des eigenen Standpunkts und auf Anfechtung der Entscheidung gehört.“ „Profiling“ [bezeichnet] jede Art der automatisierten Verarbeitung personenbezogener Daten, die darin besteht, dass diese personenbezogenen Daten verwendet werden, um bestimmte persönliche Aspekte, die sich auf eine natürliche Person beziehen, zu bewerten, insbesondere um Aspekte bezüglich Arbeitsleistung, wirtschaftliche Lage, Gesundheit, persönliche Vorlieben, Interessen, Zuverlässigkeit, Verhalten, Aufenthaltsort oder Ortswechsel dieser natürlichen Person zu analysieren oder vorherzusagen.“
Artikel 32 DSGVO Datensicherheit	„Sicherheit der Verarbeitung: Der Verantwortliche [trifft] geeignete technische und organisatorische Maßnahmen, um ein dem Risiko angemessenes Schutzniveau zu gewährleisten; diese Maßnahmen schließen unter anderem Folgendes ein: die Pseudonymisierung und Verschlüsselung personenbezogener Daten; die Fähigkeit, die Vertraulichkeit, Integrität, Verfügbarkeit und Belastbarkeit der Systeme [...] sicherzustellen; ein Verfahren zur regelmäßigen Überprüfung, Bewertung und Evaluierung der Wirksamkeit der technischen und organisatorischen Maßnahmen zur Gewährleistung der Sicherheit der Verarbeitung.“



In welchen Rechtsrahmen ist der Einsatz generativer KI eingebettet?

Artikel 35 DSGVO Datenschutzfolgen- abschätzung

Eine Folgenabschätzung ist insbesondere in folgenden Fällen erforderlich:

„[die] Verwendung neuer Technologien aufgrund der Art, des Umfangs, der Umstände und der Zwecke der Verarbeitung voraussichtlich ein hohes Risiko für die Rechte und Freiheiten natürlicher Personen [zur Folge hat];

[...] Profiling [...] als Grundlage für Entscheidungen [dient], die Rechtswirkung gegenüber natürlichen Personen entfalten oder diese in ähnlich erheblicher Weise beeinträchtigen;

[eine] umfangreiche Verarbeitung besonderer Kategorien von personenbezogenen Daten gem. Artikel 9 Absatz 1 [...]“ stattfindet.

Betriebsverfassungsgesetz

Das Betriebsverfassungsgesetz ist die Arbeitsgrundlage für alle Betriebsräte in Deutschland, die Personalvertretungsgesetze die adäquaten und oft textgleichen Gesetze für den öffentlichen Dienst. Hier werden exemplarisch jene Paragraphen des BetrVG vorgestellt, die für die Einflussnahme auf KI-Systeme bedeutsam sein können.

Beispielsweise gibt das Gesetz auch dem Arbeitgeber die Pflicht vor, diskriminierende Softwareentscheidungen zu verhindern und die freie Entfaltung der Persönlichkeit im Betrieb zu fördern. Der Betriebsrat kann Sachverständige für seine Aufgaben im Zusammenhang mit der Beurteilung von KI-Systemen hinzuziehen und hat umfassende Mitbestimmungsrechte bei Gestaltung der Ordnung im Betrieb, bei der Einführung neuer technischer Einrichtungen und beim Arbeits- und Gesundheitsschutz. Selbst der betriebliche Umweltschutz steht auf der Liste jener Aufgaben, die für den Betriebsrat wesentlich sind. Geht es um qualitative Personalplanung, kann der Betriebsrat mitreden. Ebenso bei Fragen der Beschäftigungssicherheit. Die nachfolgende Aufstellung skizziert lediglich die wichtigsten Paragraphen, um sie in Erinnerung zu rufen. Die vielfältigen Möglichkeiten, die sich aus den Paragraphen des Betriebsverfassungsgesetzes ergeben, wird den Mitbestimmungsträgern auf grundlegenden Seminaren vermittelt.

§ 75 BetrVG Behandlung von Betriebs- angehörigen

Arbeitgeber und Betriebsrat haben

- darüber zu wachen, dass alle im Betrieb tätigen Personen nach den Grundsätzen von Recht und Billigkeit behandelt werden, insbesondere, dass jede (diskriminierende) **Benachteiligung** von Personen **unterbleibt**.
- die **freie Entfaltung der Persönlichkeit** der im Betrieb beschäftigten Arbeitnehmer zu schützen und zu fördern.
- die **Selbständigkeit und Eigeninitiative** der Arbeitnehmer und Arbeitsgruppen zu fördern.

§ 80 Abs 3 BetrVG

Muss der Betriebsrat zur Durchführung seiner Aufgaben die Einführung oder Anwendung von **Künstlicher Intelligenz** beurteilen, gilt insoweit die Hinzuziehung eines **Sachverständigen** als erforderlich.

§ 87 Abs. 1 BetrVG Mitbestimmungsrechte	Der Betriebsrat hat, soweit eine gesetzliche oder tarifliche Regelung nicht besteht, darüber mitzubestimmen: [...] - Fragen der Ordnung des Betriebs und des Verhaltens der Arbeitnehmer im Betrieb (§ 87 Abs. 1 Nr. 1 BetrVG). - Einführung und Anwendung von technischen Einrichtungen, die dazu bestimmt sind, das Verhalten oder die Leistungen der Arbeitnehmer zu überwachen (§ 87 Abs. 1 Nr. 6 BetrVG). - Arbeits- und Gesundheitsschutz: Regelungen über die Verhütung von Arbeitsunfällen und Berufskrankheiten sowie über den Gesundheitsschutz im Rahmen der gesetzlichen Vorschriften oder der Unfallverhütungsvorschriften (§ 87 Abs. 1 Nr. 7 BetrVG). - „Telearbeit“: Ausgestaltung von mobiler Arbeit, die mittels Informations- und Kommunikationstechnik erbracht wird (§ 87 Abs. 1 Nr. 14 BetrVG).
§ 89 BetrVG Arbeits- und betrieblicher Umweltschutz	Der Betriebsrat hat sich dafür einzusetzen, dass die Vorschriften über den Arbeitsschutz und die Unfallverhütung im Betrieb sowie über den betrieblichen Umweltschutz durchgeführt werden.
§ 90 BetrVG Unterrichtungs- und Beratungsrechte	(1) Der Arbeitgeber hat den Betriebsrat u.a. über die Planung - von technischen Anlagen, - von Arbeitsverfahren und Arbeitsabläufen einschließlich des Einsatzes von Künstlicher Intelligenz rechtzeitig unter Vorlage der erforderlichen Unterlagen zu unterrichten.
§ 91 BetrVG Mitbestimmungsrecht	Werden die Arbeitnehmer durch Änderungen der Arbeitsplätze, des Arbeitsablaufs oder der Arbeitsumgebung, die den gesicherten arbeitswissenschaftlichen Erkenntnissen über die menschengerechte Gestaltung der Arbeit offensichtlich widersprechen, in besonderer Weise belastet, so kann der Betriebsrat angemessene Maßnahmen zur Abwendung, Milderung oder zum Ausgleich der Belastung verlangen.
§ 92 BetrVG Personalplanung	Der Arbeitgeber hat den Betriebsrat über die Personalplanung, insbesondere über den gegenwärtigen und künftigen Personalbedarf sowie über die sich daraus ergebenden personellen Maßnahmen einschließlich der geplanten Beschäftigung von Personen, die nicht in einem Arbeitsverhältnis zum Arbeitgeber stehen, und Maßnahmen der Berufsbildung anhand von Unterlagen rechtzeitig und umfassend zu unterrichten.
§ 92a BetrVG Beschäftigungssicherung	„(1) Der Betriebsrat kann dem Arbeitgeber Vorschläge zur Sicherung und Förderung der Beschäftigung machen. Diese können insbesondere [...] neue Formen der Arbeitsorganisation, Änderungen der Arbeitsverfahren und Arbeitsabläufe, die Qualifizierung der Arbeitnehmer [...] zum Gegenstand haben.“
§ 95 BetrVG Auswahlrichtlinien	„1) Richtlinien über die personelle Auswahl bei Einstellungen, Versetzungen, Umgruppierungen und Kündigungen bedürfen der Zustimmung des Betriebsrats. [...] (2) In Betrieben mit mehr als 500 Arbeitnehmern kann der Betriebsrat die Aufstellung von Richtlinien, über die bei Maßnahmen des Absatzes 1 Satz 1 zu beachtenden fachlichen und persönlichen Voraussetzungen und sozialen Gesichtspunkte verlangen.“



Hinweisgeberschutzgesetz

Seit Mitte 2023 existiert in Deutschland ein Hinweisgeberschutzgesetz. Die Bundesregierung hatte zwar eine Strafzahlung zu akzeptieren, weil sie ein entsprechendes EU-Gesetz erst verspätet in Deutschland umgesetzt hatte. Jetzt aber ist ein Gesetz gültig, dass äußerst hilfreich dabei ist, die Betroffenen zu Beteiligten zu machen. Beschäftigte können sich an die internen oder externen Meldestellen wenden, wenn sie Rechtsvorschriften verletzt sehen oder Indizien dafür haben, dass ihre personenbezogenen Daten und ihre Privatsphäre nicht ausreichend geschützt sind. Sie dürfen wegen einer entsprechenden Meldung keine Repressalien erfahren. Der regelgerechte Einsatz generativer KI kann mit den Möglichkeiten dieses Gesetzes überwacht werden.

§2 Sachlicher Anwendungsbereich	„Dieses Gesetz gilt für die Meldung [...] und die Offenlegung [...] von Informationen über: - Verstöße, die strafbewehrt sind, - Verstöße, die bußgeldbewehrt sind, soweit die verletzte Vorschrift dem Schutz von Leben, Leib oder Gesundheit oder dem Schutz der Rechte von Beschäftigten oder ihrer Vertretungsorgane dient, - sonstige Verstöße gegen Rechtsvorschriften des Bundes und der Länder sowie unmittelbar geltende Rechtsakte der Europäischen Union [...] o) zum Schutz der Privatsphäre [...] p) zum Schutz personenbezogener Daten (DSGVO) [...] q) zur Sicherheit in der Informationstechnik im Sinne des § 2 Absatz 2 des BSI-Gesetzes.“
§7 Wahlrecht zwischen interner und externer Meldung	„(1) Personen, die beabsichtigen, Informationen über einen Verstoß zu melden, können wählen, ob sie sich an eine interne Meldestelle (§ 12) oder eine externe Meldestelle (§§ 19 bis 24) wenden. Wenn einem intern gemeldeten Verstoß nicht abgeholfen wurde, bleibt es der hinweisgebenden Person unbenommen, sich an eine externe Meldestelle zu wenden. (2) Es ist verboten, Meldungen oder die auf eine Meldung folgende Kommunikation zwischen hinweisgebender Person und Meldestelle zu behindern oder dies zu versuchen.“
§ 12 Pflicht zur Einrichtung interner Meldestellen	„(1) Beschäftigungsgeber haben dafür zu sorgen, dass bei ihnen mindestens eine Stelle für interne Meldungen eingerichtet ist und betrieben wird, an die sich Beschäftigte wenden können (interne Meldestelle).“
§36 Verbot von Repressalien; Beweislastumkehr	„(1) Gegen hinweisgebende Personen gerichtete Repressalien sind verboten. Das gilt auch für die Androhung und den Versuch, Repressalien auszuüben. (2) Erleidet eine hinweisgebende Person nach einer Meldung oder Offenlegung eine Benachteiligung im Zusammenhang mit ihrer beruflichen Tätigkeit, so wird vermutet, dass diese Benachteiligung eine Repressalie ist. In diesem Fall hat die Person, die die hinweisgebende Person benachteiligt hat, zu beweisen, dass die Benachteiligung auf hinreichend gerechtfertigten Gründen basierte oder dass sie nicht auf der Meldung oder Offenlegung beruhte.“

Allgemeines Gleichbehandlungsgesetz

Diskriminierungsfreiheit ist ein gern formulierter Anspruch im Rahmen der Gestaltung von KI-Systemen. In der innerbetrieblichen Wirklichkeit erhält er aber häufig nicht die Aufmerksamkeit, die er verdient. Werden Betriebsräte und Personalräte danach gefragt, an welchen Faktoren sie ihr Vertrauen in KI-Technologie festmachen, wird der Vorrang menschlichen Handelns, der Schutz personenbezogener Daten, die Aufsicht und Kontrolle über derartige Systeme zuerst genannt. Nichtdiskriminierung und Fairness rangieren eher auf den unteren Plätzen. Dies ist Anlass genug, an die Vorgaben des Allgemeinen Gleichbehandlungsgesetzes zu erinnern.

§ 1 AGG Ziel des Gesetzes	„ Benachteiligungen aus Gründen der Rasse oder wegen der ethnischen Herkunft, des Geschlechts, der Religion oder Weltanschauung, einer Behinderung, des Alters oder der sexuellen Identität [gilt es] zu verhindern oder zu beseitigen. “
§ 2 AGG Anwendungsbereich	„Benachteiligungen [...] sind unzulässig in Bezug auf: [...] Auswahlkriterien und Einstellungsbedingungen , [...] den beruflichen Aufstieg , die Beschäftigungs- und Arbeitsbedingungen , [die] Berufsbildung, [und die] praktische Berufserfahrung.“
§12 AGG Maßnahmen und Pflichten des Arbeitgebers	Der Arbeitgeber ist verpflichtet, die erforderlichen Maßnahmen zum Schutz vor Benachteiligungen wegen eines in § 1 genannten Grundes zu treffen. Dieser Schutz umfasst auch vorbeugende Maßnahmen .

KI-Verordnung der EU

Anfang 2024 wurde endlich die Europäische KI-Verordnung vom EU-Parlament verabschiedet und damit ein großer Meilenstein für die Beachtung von Grundrechten bei der Nutzung von künstlicher Intelligenz gelegt. Die Europäische KI-Verordnung wirkt sich unmittelbar auf die Mitgliedsstaaten aus und die Umsetzung in den Betrieben muss nun vorangebracht werden. Betriebs- und Personalräte haben mit der neuen Rechtsvorschrift eine gute Grundlage, auf die Einführung von KI-Systemen im Betrieb Einfluss zu nehmen.

Die Europäische KI-Verordnung unterscheidet zwischen KI-Anwendungen mit geringem Risiko und KI-Systemen mit hohem Gefahrenpotenzial. Für letztere sind besondere Auflagen zu erfüllen. Ist das Risiko hoch, dann werden auch hohe Anforderungen an das KI-System gesetzt. Sehr kritische Systeme sind künftig in Europa verboten.

Die EU-Verordnung stärkt auch einen Rechtsanspruch, den die deutsche Mitbestimmung in ähnlicher Weise bereits aus § 90 Abs. 1 Nr. 3 Betriebsverfassungsgesetz kennt:

„Vor der Inbetriebnahme oder Verwendung eines Hochrisiko-KI-Systems am Arbeitsplatz informieren die Betreiber, die Arbeitgeber sind, die Arbeitnehmervertreter und die betroffenen Arbeitnehmer darüber, dass sie dem System unterworfen sein werden.“^{57 58}

Rechtzeitige Information der Arbeitnehmervertretungen ist die Voraussetzung dafür, um überhaupt Schutzmechanismen einfordern zu können. Darauf zielt die Europäische Kommission nun ab und sie geht noch weiter: Wer nicht vor der Inbetriebnahme des KI-Systems Arbeitnehmer-

57 Dieser Artikel nimmt Bezug auf den Text der Europäischen KI-Verordnung, angenommen vom Europäischen Parlament am 13. März 2024.

58 Art. 26 Abs. 7 EU KI-VO.



vertretungen informiert, läuft Gefahr, Verwaltungsstrafen bis zur Höhe von 15 Millionen Euro oder 3 Prozent des gesamten weltweiten Jahresumsatzes des Unternehmens zu riskieren.⁵⁹ Aber zunächst ist zu klären, was ein KI-System eigentlich ist.

Was ist ein KI-System?

Erstmals hat der europäische Gesetzgeber eine Definition für KI-Systeme vorgenommen. Hier nach gilt künftig:

Ein KI-System ist „ein maschinengestütztes System, das für einen in wechselndem Maße autonomen Betrieb ausgelegt ist, das nach seiner Einführung anpassungsfähig sein kann und das aus seinen Eingaben für explizite oder implizite Ziele ableitet, wie Ergebnisse, wie etwa Vorhersagen, Inhalte, Empfehlungen oder Entscheidungen hervorgebracht werden, die physische oder virtuelle Umgebungen beeinflussen können“.⁶⁰

Es dürfte trotz dieser definitorischen Klarstellung nicht leichtfallen, in der betrieblichen Praxis zu klären, wo herkömmliche IT-Systeme aufhören und wo KI-Systeme anfangen.

Unterstützung findet man dabei in der Kommentierung von Fitting u.a. zum Betriebsverfassungsgesetz: „Charakteristisch für KI ist- in Abgrenzung zu jeglicher Informations- und Kommunikationstechnik- die mangelnde Vorhersehbarkeit der zu findenden Ergebnisse“.⁶¹ Hilfreich zur Abgrenzung sind auch die Leitlinien der Kommission zur Definition eines Systems der Künstlichen Intelligenz gemäß der Verordnung (EU) 2024/1689 (KI-Gesetz), die die EU-Kommission im Mai 2025 herausgegeben hat.

Für Betriebs- und Personalräte ist es zunächst notwendig zu klären, ob sie es überhaupt mit einem KI-System zu tun haben. Weil die EU-Definition tatsächlich nicht selbsterklärend ist, hat der Konzernbetriebsrat der Provinzialversicherung AG mit einer Regelungsabrede u. a. eine Clearingstelle eingerichtet, die innerbetrieblich die Frage klären soll, wo IT aufhört und KI anfängt. Das kann betriebliche, juristische und technische Kontroversen eindämmen.

Transparenzvorgaben⁶²

Die Europäische KI-Verordnung macht für die Bereitstellung bestimmter Anwendungen, insbesondere von generativer KI, zahlreiche Transparenzverpflichtungen. So ist es jetzt zur Vorgabe geworden, dass Beschäftigte darüber informiert werden müssen, wenn sie mit einem KI-System interagieren. Anbieter von generativer KI, wie z.B. ChatGPT „**die synthetische Audio-, Bild-, Video-, oder Textinhalte erzeugen**“, müssen sicherstellen, dass „**die Ausgaben des KI-Systems in einem maschinenlesbaren Format gekennzeichnet sind und als künstlich erzeugt und manipuliert erkannt werden können**.“

Die Betreiber eines Systems zur biometrischen Kategorisierung müssen die Personen, die solchen Systemen ausgesetzt sind, über den Betrieb eines solchen Systems informieren. Der Einsatz von Deep Fakes muss offengelegt werden und es muss darüber in Kenntnis gesetzt werden, dass „**die Inhalte künstlich erzeugt oder manipuliert wurden**“. Auch die Pflicht zur Bereitstellung technischer Informationen darf als weiteres Instrument begriffen werden, Transparenz über die Funktionsmechanismen von KI-Systemen zu schaffen.

59 Art. 99 Abs. 4 EU KI-VO.

60 Art. 3 Abs. 1 EU KI-VO.

61 RN 94 ff zu § 80 BetrVG, Fitting u.a.: Betriebsverfassungsgesetz, Handkommentar, 32.Auflage 2024.

62 Insbesondere EU KI VO Artikel 50.

Verbotene Systeme

In allen Mitgliedsstaaten der Europäischen Union und damit auch in deutschen Betrieben wird mit der KI-Verordnung u.a. der Einsatz von KI-Systemen verboten⁶³, die

- das Verhalten von Personen mit unterschwelligem Techniken außerhalb des Bewusstseins zur gezielten Manipulation einsetzen;
- Schwächen einer Person oder einer bestimmten Personengruppe aufgrund ihres Alters, einer Behinderung oder einer besonderen sozialen oder wirtschaftlichen Situation ausnutzen;
- biometrische Kategorisierung nutzen, um Rückschlüsse auf die Rasse, die politischen Meinungen, die Gewerkschaftszugehörigkeit, die religiösen oder philosophischen Überzeugungen, das Sexualleben oder sexuelle Ausrichtung ziehen oder zulassen;
- Menschen aufgrund ihres Sozialverhaltens oder persönlicher Merkmale klassifizieren und benachteiligen;
- am Arbeitsplatz oder in Bildungseinrichtungen Emotionserkennung durchführen, die über medizinische Zwecke oder Sicherheitsanwendungen hinausgehen.

Verboten ist zudem die Bewertung von sozialem Verhalten (social scoring), wie man es aus China kennt.

KI-Systeme mit hohem Risiko

Die Europäische KI-Verordnung ordnet auch zahlreiche Anwendungen dem Hochrisikobereich zu, von denen Beschäftigte betroffen sind. Im Anhang III der Europäischen KI-Verordnung werden Anwendungen des Hochrisikobereichs benannt und dieser Anhang soll künftig fortentwickelt werden. Zum Zeitpunkt des Inkrafttretens der Verordnung zählen erstmalig u.a. folgende KI-Systeme zum Hochrisikobereich:

- „KI-Systeme, die für die Einstellung oder Auswahl natürlicher Personen verwendet werden sollen, insbesondere um gezielte Stellenanzeigen zu schalten, Bewerbungen zu sichten und zu filtern und Bewerber zu bewerten;
- KI, die bestimmungsgemäß für Entscheidungen, die die Bedingungen von Arbeitsverhältnissen, Beförderungen und Kündigungen von Arbeitsvertragsverhältnissen beeinflussen, für die Zuweisung von Aufgaben auf der Grundlage des individuellen Verhaltens oder persönlicher Merkmale oder Eigenschaften und Bewertung der Leistung und des Verhaltens von Personen in solchen Beschäftigungsverhältnissen verwendet werden sollen;
- KI-Systeme, die bestimmungsgemäß für die biometrische Kategorisierung nach sensitiven oder geschützten Attributen oder Merkmalen oder auf der Grundlage von Rückschlüssen auf diese Attribute oder Merkmale verwendet werden sollen;
- KI-Systeme, die bestimmungsgemäß als Sicherheitskomponenten im Rahmen der Verwaltung und des Betriebs kritischer digitaler Infrastruktur, des Straßenverkehrs sowie der Wasser-, Gas-, Wärme- und Stromversorgung verwendet werden sollen;
- KI-Systeme, die bestimmungsgemäß zur Feststellung des Zugangs oder der Zulassung oder zur Zuweisung natürlicher Personen zu Einrichtungen aller Ebenen der allgemeinen und beruflichen Bildung verwendet werden sollen;
- KI-Systeme, die bestimmungsgemäß für die Bewertung von Lernergebnissen verwendet werden sollen, einschließlich des Falles, dass diese Ergebnisse dazu dienen, den Lernprozess natürlicher Personen in Einrichtungen oder Programmen aller Ebenen der allgemeinen und beruflichen Bildung zu steuern;

63 EU KI-VO Art. 3.



- KI-Systeme, die bestimmungsgemäß zum Zweck der Bewertung des angemessenen Bildungsniveaus, das eine Person im Rahmen von oder innerhalb von Einrichtungen der allgemeinen und beruflichen Bildung erhalten wird oder zu dem sie Zugang erhalten wird, verwendet werden sollen;
- KI-Systeme, die bestimmungsgemäß zur Überwachung und Erkennung von verbotenen Verhalten von Schülern bei Prüfungen im Rahmen von oder innerhalb von Einrichtungen der allgemeinen und beruflichen Bildung verwendet werden sollen.“⁶⁴

Auch ein Profiling von Beschäftigten, gilt in den Anwendungsfällen, die im Anhang III beschrieben sind, immer als mit hohem Risiko behaftet.⁶⁵

Systeme dagegen, die kein erhebliches Risiko für die Gesundheit, die Sicherheit oder die Grundrechte darstellen gelten nicht als hochrisikobehaftet. Dazu werden Bedingungen genannt:

- das KI-System ist dazu bestimmt, eine eng gefasste Verfahrensaufgabe durchzuführen;
- das KI-System ist dazu bestimmt, das Ergebnis einer zuvor abgeschlossenen menschlichen Tätigkeit zu verbessern;
- das KI-System ist dazu bestimmt, Entscheidungsmuster oder Abweichungen von früheren Entscheidungsmustern zu erkennen, und ist nicht dazu gedacht, die zuvor abgeschlossene menschliche Bewertung, ohne eine angemessene menschliche Überprüfung zu ersetzen oder zu beeinflussen; oder
- das KI-System ist dazu bestimmt, eine vorbereitende Aufgabe für eine Bewertung durchzuführen, die für die Zwecke, der in Anhang III aufgeführten Anwendungsfälle relevant ist.⁶⁶

Grundrechtsfolgenabschätzung

Die Wirkungen auf die Gesundheit, die Sicherheit oder die Grundrechte festzustellen oder die Wirkungstiefe eines Systems zu beurteilen, wird ohne eine Grundrechtsfolgeabschätzung, kaum möglich sein. Diese wiederum gibt die KI-Verordnung für Hochrisikosysteme vor⁶⁷ – zumindest für KI-Betreiber in Einrichtungen öffentlichen Rechts oder private Betreiber, die in deren Auftrag handeln. Sie muss vor dem Einsatz eines KI-Systems mit hohem Risiko durchgeführt werden und u.a. folgende Informationen umfassen:

- die Kategorien der natürlichen Personen und Personengruppen, die von seiner Verwendung im spezifischen Kontext betroffen sein könnten;
- die spezifischen Schadensrisiken, die sich auf die [...] ermittelten Kategorien von Personen oder Personengruppen auswirken könnten, unter Berücksichtigung der vom Anbieter gemäß Artikel 13 bereitgestellten Informationen;
- eine Beschreibung der Umsetzung von Maßnahmen der menschlichen Aufsicht entsprechend den Gebrauchsanweisungen;
- die Maßnahmen, die bei Eintreten dieser Risiken zu ergreifen sind, einschließlich der Regelungen für die interne Unternehmensführung und Beschwerdemechanismen.

Betriebs und Personalräte sind gut beraten, wenn sie sich vor der Einführung eines KI-Systems, das hoch risikobehaftet ist, die erforderliche Grundrechtsfolgenabschätzung vorlegen lassen. Noch besser wäre es, wenn sie darauf Einfluss nähmen, mit welcher Methode und mit welchen

64 EU VO Anlage III.

65 EU VO Art. 6 Abs 3.

66 EU KI-VO Artikel 6 Absatz 3.

67 EU VO Art 27.

weiteren Untersuchungsgegenständen eine Grundrechtsfolgeabschätzung im Betrieb durchgeführt wird. Die wäre in allen Betrieben erforderlich, schließlich haben alle Beschäftigte Grundrechte und es dürfte mit der Grundrechtscharta kaum verträglich sein, wenn sich der Arbeitsdruck erhöht, Überwachung zunimmt oder im erheblichem Umfang Arbeitsplätze abgebaut werden.

Schließlich sind sichere, gesunde und würdige Arbeitsbedingungen schon ein Anspruch der Grundrechtscharta, dem auch die Europäische KI-Verordnung verpflichtet ist.⁶⁸

Im deutschen Arbeitsschutzrecht ist dafür auch eine Gefährdungsbeurteilung vorgegeben, die ebenfalls eine vorausschauende Beurteilung der Folgen des Technikeinsatzes vorsieht, nicht nur für hochriskante Systeme.

Auflagen für Hochrisikosysteme⁶⁹

Für Hochrisikosysteme gibt die Europäische KI-Verordnung erhebliche Anforderungen vor, die Unternehmen und Verwaltungen einzuhalten haben. Dazu zählt ein kontinuierliches, iteratives Risikomanagement über den gesamten Lebenszyklus des KI-Systems. Dabei ist die Bearbeitung der Systemrisiken nicht nur bei der Einführung einer Anwendung erforderlich. KI-Systeme mit hohem Risiko müssen laufend getestet und während der Tests anhand von festgelegten Messgrößen und Wahrscheinlichkeitsschwellen evaluiert werden.

Es muss überprüft werden, ob die Systeme für den beabsichtigten Zweck konsistent funktionieren; Trainings-, Validierungs- und Testdatensätze müssen geeignet sein. Dies ist bedeutsam, um beispielsweise Diskriminierungsgefahren einzudämmen. Deswegen muss die Herkunft der Daten untersucht werden sowie ihre Verfügbarkeit, Menge und Eignung für den eingesetzten Zweck untersucht und geprüft werden, ob Gefahren für mögliche Verzerrungen sich negativ auf die Grundrechte auswirken. Für Hochrisikosysteme müssen sehr umfassende technische Dokumentationen angelegt werden. Der zu erwartende Output des Systems muss detailliert beschrieben sein, ebenso die menschlichen Aufsichtsmaßnahmen, die für den Einsatz des Systems vorgesehen sind. Zu den technischen Unterlagen muss eine Kopie der EU-Konformitätserklärung beigefügt sein. Daraus lässt sich erkennen, ob die Systeme mit den Vorgaben der EU in Einklang stehen.

Eine Protokollierungsfunktion (Log-Datei) muss realisiert werden. Damit entstehen beim Einsatz von Hochrisikosystemen immer personenbezogene Daten, die Mitbestimmungsrechte auslösen. Die EU-Vorgaben verlangen auch danach, dass es für das jeweilige System eine Gebrauchsanweisung gibt. Diese muss *„präzise, vollständige, korrekte und eindeutige Informationen enthalten, die für die Nutzer relevant, barrierefrei zugänglich und verständlich sind.“* Die Gebrauchsanweisung muss Merkmale, Fähigkeiten und Leistungsgrenzen des Hochrisikosystems beschreiben, den vorgesehenen Zweck benennen, das Genauigkeitsniveau, mit dem das System arbeitet, transparent machen und jene Umstände beschreiben, die Grundrechte gefährden können, wie etwa Fehlanwendungen des Systems. Eine derartige Gebrauchsanweisung durchzusehen, kann die Meinungsbildung von Mitbestimmungsträger*innen oder deren Sachverständige erheblich erleichtern.

68 EU Grundrechtscharta Art 31.

69 EU VO Art 8 bis 10.



Menschliche Aufsicht⁷⁰

Menschliche Aufsicht wird zum Betrieb von Hochrisiko-Systemen vorgegeben. Damit sollen *„Risiken für die Gesundheit, Sicherheit oder die Grundrechte verhindert oder auf ein Mindestmaß reduziert werden.“* Jene Personen, die Aufsicht führen, sollen in die Lage versetzt werden:

- „die einschlägigen Fähigkeiten und Grenzen des Hochrisiko-KI-Systems angemessen zu verstehen und seinen Betrieb ordnungsgemäß zu überwachen, auch in Bezug auf das Erkennen und Beheben von Anomalien, Fehlfunktionen und unerwarteter Leistung;
- sich einer möglichen Neigung zu einem automatischen oder übermäßigen Vertrauen in das von einem Hochrisiko-KI-System hervorgebrachte Ergebnis („Automatisierungsbias“) bewusst zu bleiben, insbesondere wenn Hochrisiko-KI-Systeme Informationen oder Empfehlungen ausgeben, auf deren Grundlage natürliche Personen Entscheidungen treffen;
- das Ergebnis des Hochrisiko-KI-Systems richtig zu interpretieren, wobei beispielsweise die vorhandenen Interpretationsinstrumente und -methoden zu berücksichtigen sind;
- in einer bestimmten Situation zu beschließen, das Hochrisiko-KI-System nicht zu verwenden oder das Ergebnis des Hochrisiko-KI-Systems außer Acht zu lassen, außer Kraft zu setzen oder rückgängig zu machen;
- in den Betrieb des Hochrisiko-KI-Systems einzugreifen oder den Systembetrieb mit einer „Stopptaste“ oder einem ähnlichen Verfahren zu unterbrechen, was dem System ermöglicht, in einem sicheren Zustand zum Stillstand zu kommen.“

Kompetenz⁷¹

Die Europäische KI-Verordnung verlangt von den Arbeitgebern, für alle KI-Systeme die für den Betrieb der KI erforderliche Kompetenz aufzubauen:

„Die Anbieter und Betreiber von KI-Systemen ergreifen Maßnahmen, um nach besten Kräften sicherzustellen, dass ihr Personal und andere Personen, die in ihrem Auftrag mit dem Betrieb und der Nutzung von KI-Systemen befasst sind, über ein ausreichendes Maß an KI-Kompetenz verfügen, wobei ihre technischen Kenntnisse, ihre Erfahrung, ihre Ausbildung und Schulung und der Kontext, in dem die KI-Systeme eingesetzt werden sollen, sowie die Personen oder Personengruppen, bei denen die KI-Systeme eingesetzt werden sollen, zu berücksichtigen sind.“

Auch mit dieser neuen Vorgabe lässt sich eine Anforderung von Betriebs- und Personalräten auf eine seriöse qualitative Personalplanung begründen.

70 EU VO Art. 14.

71 EU VO Art. 4.

Marktaufsichtsbehörden⁷²

Künftig werden nicht nur betriebliche Mitbestimmungsträger*innen über die Einhaltung rechtlicher Vorgaben und ethischer Standards wachen. Jeder Mitgliedstaat der Europäischen Union hat künftig eine Marktaufsichtsbehörde einzurichten, die die Einhaltung der KI-Verordnung überwacht. Die Verordnung gibt allen Bürgern Europas das Recht, die Marktüberwachungsbehörde anzurufen, wenn sie Grund zu der Annahme haben, dass ein Verstoß gegen die Bestimmungen der Verordnung vorliegt.⁷³ Damit ergeben sich auch für die Beschäftigten im Betrieb oder in den Mitbestimmungsgremien zusätzliche Eskalationsmöglichkeiten, sollte ein Betrieb die Anwendung der KI-Verordnung nicht allzu ernst nehmen.

Freiwillige Verhaltenskodizes und Partizipation⁷⁴

Die Europäische Kommission will Verhaltenskodizes fördern, die für KI-Systeme ohne hohes Risiko gelten. Das dürfte die Verabredung von Betriebs- und Dienstvereinbarungen erleichtern, die Prinzipien für den KI-Einsatz vorgeben und Prozesse der Mitbestimmung regeln. Die Kommission betont, dass solche Verhaltenskodizes unter Einbeziehung relevanter Interessensvertretungen entwickelt werden können und sich insbesondere an folgenden Elementen orientieren sollten:

- Anwendbare Grundsätze aus den Ethikleitlinien der EU für eine vertrauenswürdige KI;
- Bewertung und Minimierung der ökologischen Auswirkungen von KI-Systemen, insbesondere durch energieeffiziente Programmierung sowie den Einsatz von Techniken zur effizienten Gestaltung, zum Training und zur Nutzung von KI;
- Förderung der KI-Kompetenz, insbesondere bei Personen, die an der Entwicklung, dem Betrieb und der Anwendung von KI beteiligt sind;
- Unterstützung einer inklusiven und vielfältigen Gestaltung von KI-Systemen, unter anderem durch die Bildung diverser Entwicklungsteams und die aktive Einbindung von Interessenträgern in den Gestaltungsprozess;
- Bewertung und Vermeidung negativer Auswirkungen von KI-Systemen auf schutzbedürftige Personen oder Gruppen, einschließlich der Sicherstellung barrierefreier Zugänglichkeit für Menschen mit Behinderungen sowie der Förderung der Gleichstellung der Geschlechter.

72 EU VO Art. 70.

73 EU VO Art. 85.

74 EU VO Art. 95.

WAS IST BEIM
BESCHÄFTIGTEN-
DATENSCHUTZ
ZU BEACHTEN?





WAS IST BEIM BESCHÄFTIGTEN-DATENSCHUTZ ZU BEACHTEN?

BEDAX-Prüftool zur Selbstbewertung von KI-Anwendungen

Weltweit werden immer mehr KI-Systeme in Betrieben eingesetzt, um Effizienz und Effektivität zu erhöhen, neue Geschäftsmodelle zu entwickeln, Leistungs- und Prozessqualität zu verbessern, Ressourcen einzusparen, Analytik zu verbessern und Handlungsgeschwindigkeit auszubauen. KI bietet auch im gewerblichen Einsatz immenses Potenzial, aber auch erhebliche Gefahren. Zu den Gefahren gehören die mangelnde Transparenz bei KI-gesteuerten Prozessen, Kontrollverlust, das Potenzial für Diskriminierung und das Risiko der Verletzung von Persönlichkeitsrechten. Der Schutz personenbezogener Daten ist ein Persönlichkeitsrecht, das besondere Aufmerksamkeit verdient. Beim Einsatz von KI-Systemen in der Arbeitswelt sind diese so zu gestalten, dass Beschäftigte und ihre Daten geschützt, Arbeit erleichtert und die Handlungs- und Gestaltungsspielräume der Beschäftigten erweitert werden. Interessenvertretungen (BR, PersR, MAV) spielen dabei eine zentrale Rolle.

Das Projekt BeDaX⁷⁵ (Index Beschäftigtendatenschutz) hat gemeinsam mit der Schröfers IgA GmbH und mit Unterstützung von ver.di ein Prüftool zur Selbstbewertung der Datenschutzaspekte bei KI-Anwendungen erarbeitet. Es ist ein Instrument zur Unterstützung der Mitbestimmungsakteur*innen und umsichtiger betrieblicher Praktiker*innen. Untersucht werden die Wirkung von betrieblichen KI-Anwendungen sowie deren Regulierung und soziotechnische Bedingungen entsprechend den acht BeDaX Qualitätsmaßstäben, denen wiederum 28 Faktoren zugeordnet wurden. Jeder qualitätsrelevante Faktor wird mit konkreten Fragen untersucht. Damit werden die Überprüfung des Schutzniveaus für den Beschäftigtendatenschutz bei KI-Anwendungen ermöglicht und konkrete Schutzlücken aufgezeigt.

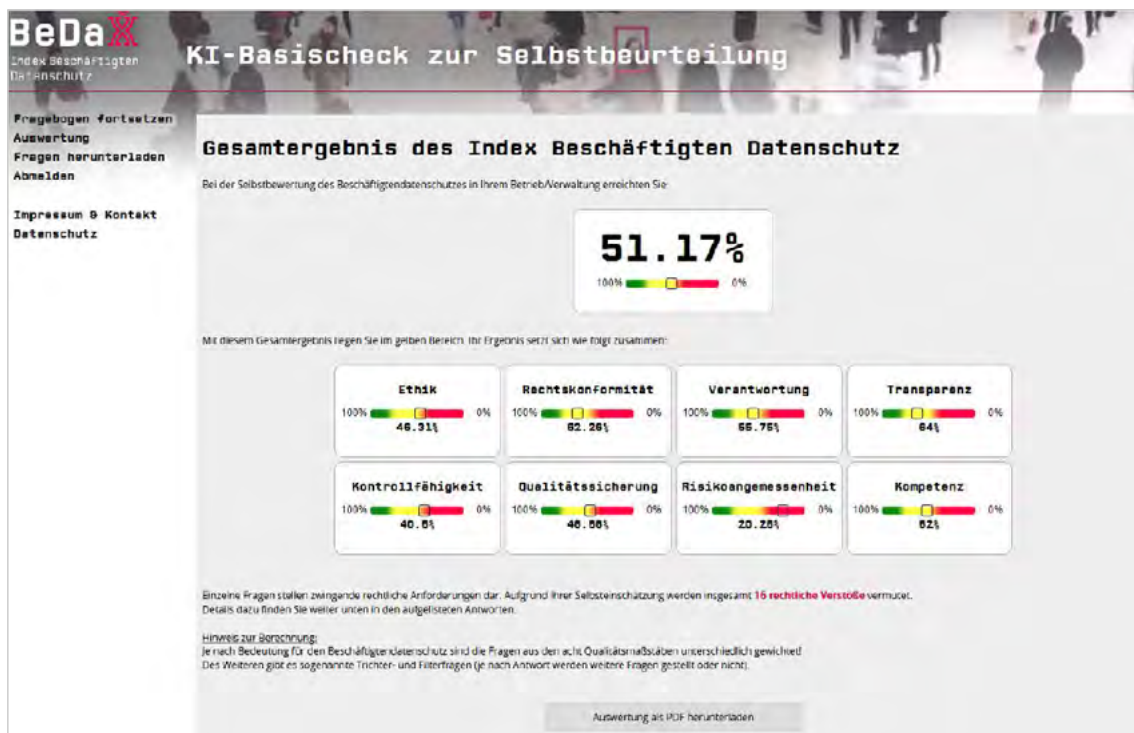
⁷⁵ <https://www.bedax.net/>

Die BeDaX-Qualitätsmaßstäbe für einen guten Beschäftigtendatenschutz

Qualitätsmaßstäbe	Faktoren
Ethik	Automatisierte Entscheidungen, Mensch steuert Maschine, Profiling, Privatsphäre, unterschwellige Verhaltensbeeinflussung, personalwirksame Schlussfolgerungen, ethische Standards
Rechtskonformität	Erforderlichkeit personenbezogener Daten, Zulässigkeit der Rechtsbasis, Zweckbindung, Diskriminierungsfreiheit und Fairness, Folgenabschätzung
Verantwortung	Datenschutzsensibler Einkauf und Entwicklung, technische Verantwortlichkeit, fachliche Verantwortlichkeit, Verantwortung für die Betroffenenrechte, Verantwortung des/der Datenschutzbeauftragten
Transparenz	Kennzeichnung und Beschreibung, Erwartungskonformität und Nachvollziehbarkeit, Dokumentation
Kontrollfähigkeit	Aufsicht, Rückholbarkeit und Interventionsmöglichkeiten, Steuerungserfahrung und Schutz vor Überinterpretation
Qualitätssicherung	Datenqualitätsmanagement, Systemtraining, Partizipation der Beschäftigten, Evaluation
Risikoangemessenheit	Risikoeinschätzung und Klassifikation, Umgang mit untragbaren und hohen Risiken
Kompetenz	Verfügbares Fachwissen und Qualifizierungsoptionen

Die Ergebnisse der Fragen werden im Erhebungstool direkt nach dem Ausfüllen der standardisierten Fragen angezeigt. Das gibt schnell wichtige Hinweise zur Ableitung von praxistauglichen, angemessenen sowie rechtskonformen Lösungen für die betriebliche Gestaltungsarbeit.

Abbildung 12: BeDAX-KI-Basischeck Gesamtergebnis



Die Fragen zu beantworten, soll aber auch dabei helfen, mehr darüber zu lernen, welche gesetzlichen Anforderungen dafür gelten, personenbezogene Daten der Beschäftigten beim Einsatz von KI-Systemen zu schützen. Deshalb ist den Fragen jeweils ein Erklärungstext vorangestellt.



Was ist beim Beschäftigtendatenschutz zu beachten?.....

Abbildung 13: BeDAX-KI-Basischeck Beispielfrage

BeDa
Index Beschäftigter
Datenschutz

KI-Basischeck zur Selbstbeurteilung

Fragebogen fortsetzen
Auswertung
Fragen herunterladen
Abmelden
Impressum & Kontakt
Datenschutz

Das Datenschutzrecht will die Würde und Selbstbestimmung der Beschäftigten wahren. Jeder und Jede hat deshalb nach der Europäischen Datenschutz-Grundverordnung das Recht, nicht einer maschinellen Entscheidung unterworfen zu werden, die ihr gegenüber rechtliche Wirkung entfaltet oder sie in ähnlicher Weise erheblich beeinträchtigt (Art. 22 DSGVO). In der Regel sind ausschließlich automatisiert getroffene Entscheidungen über Personen datenschutzrechtlich rechtswidrig. Mit einer wirksamen Einwilligung der Betroffenen sind solche Entscheidungsmethoden ausnahmsweise zulässig, aber selbst dann sind sie, gerade im Arbeitsverhältnis, an strenge Voraussetzungen gebunden. Auch eine Betriebsvereinbarung über den Einsatz derartiger Entscheidungsverfahren kann die erforderliche Einwilligung nicht ersetzen. Zu beachten ist, dass als Beschäftigte u.a. auch Bewerberinnen und Bewerber gelten.

Werden in Ihrem Betrieb oder Verwaltung Beschäftigtendaten von KI-Anwendungen verwendet, um Entscheidungen über Personen ausschließlich automatisiert zu treffen?

☒ Ja
☐ Nein

Zurück Weiter

Seite 2 von 72 | Ihr Antwort für diesen Fragebogen: KWP75VCL

© INPUT Consulting gGmbH [Seitenanfang](#)

Die spezifischen Fragen sind manchmal nicht zu beantworten, ohne vorher Informationen bei Verantwortlichen oder den Datenschutzbeauftragten zu erfragen. Das ist Absicht: Es soll bei den Teilnehmer*innen an dem Selbsttest ein umfassendes Bild zum Beschäftigtendatenschutz bei KI-Anwendungen entstehen. Auch sollen wesentliche Elemente des Themas beleuchtet werden, die eine Schlüsselfunktion haben, die aber bisweilen zu kurz kommen, weil manche Anwendungen sehr komplex sind, nicht verständlicher erläutert werden oder die Verhältnisse in manchen Betrieben es nicht erlauben. Weil Rückfragen Zeit brauchen, kann der Selbsttest jederzeit unterbrochen und später weiterbearbeitet werden.

Der Selbsttest gibt auch sogenannte Warnhinweise aus, falls Antworten darauf hindeuten, dass Kernelemente des Datenschutzes nicht erfüllt werden und dies möglicherweise sehr teure Sanktionen zur Folge haben kann. Hier ein Beispiel:

Abbildung 14: BeDAX-KI-Basischeck Beispiel für Warnhinweis

Frage 015:

Werden im Betrieb KI-Systeme eingesetzt, mit denen Schlussfolgerungen für die Entwicklung oder Nutzung der Qualifikation von Beschäftigten gezogen werden?

Antwort: Ja (10%)

Warnhinweis:
Vorsicht: KI-Systeme, mit denen Schlussfolgerungen für die Karrieren von Bewerberinnen oder Beschäftigten gezogen werden, sind nach der europäischen KI-Verordnung Hochrisikoanwendungen. Wenn diese in Ihrem Betrieb eingesetzt, müssen zumindest die hohen Regulierungsanforderungen erfüllt werden, die der Verordnungsentwurf vorsieht. Wir empfehlen einen Blick in den Gesetzesentwurf.

Gewichtung der Antworten im Selbsttest

Bei der Gewichtung der einzelnen Qualitätsmaßstäbe zu Datenschutz und Künstlicher Intelligenz ist neben der **Rechtskonformität** ein **zentraler Maßstab die Ethik**. Die Kernfrage ist, ob die Maschine den Menschen oder der Mensch die Maschine steuert. Eine wichtige Grundlage zur Einflussnahme auf die ethischen und rechtlichen Aspekte stellt die **Transparenz** dar. Die Einschätzung, inwieweit das KI-System **verständlich** und die maschinellen Schlussfolgerungen **nachvollziehbar** sind, ist bedeutsam für menschliches Vertrauen. Da KI-Systeme in ihrer Entscheidungsfindung nicht immer fehlerfrei arbeiten, stellt die **Kontrollfähigkeit** einen weiteren wichtigen Faktor dar. Kontrollfähigkeit korrespondiert mit der **Risikoeinschätzung**. Die Untersuchung von Aspekten der **Verantwortung**, der **Kompetenz** und der **Qualitätssicherung** ergänzen und vervollständigen den Basischeck.

Die Gewichtung der Antworten auf die Fragen zu den Qualitätsmaßstäben:

	Qualitätsmaßstäbe	%
1	Ethik	15
2	Rechtskonformität	25
3	Verantwortung	8
4	Transparenz	16
5	Kontrollfähigkeit	10
6	Qualitätssicherung	8
7	Risikoangemessenheit	10
8	Kompetenz	8
	Summe	100

Zugang zum Selbsttest

Wer Interesse daran hat, einen KI-Datenschutz-Selbsttest zu machen, kann per E-Mail (info@bedax.net) einen Zugangscodes anfordern. Alle Informationen dazu finden Sie auch unter www.bedax.net.

Die Website www.bedax.net bietet auch ein umfangreiches Informationsangebot zum Beschäftigtendatenschutz, das laufend ausgebaut und erweitert wird. Es richtet sich an Mitbestimmungsakteur*innen und vermittelt neben Grundlagen des Beschäftigtendatenschutzes auch aktuelle Rechtsentwicklungen, Empfehlungen der Aufsichtsbehörden, nützliche Links und hilfreiche Praxisbeispiele.

WER WILL NOCH
GENERATIVE KI
MIT WELCHEN MITTELN
BEGRENZEN?





WER WILL NOCH GENERATIVE KI MIT WELCHEN MITTELN BEGRENZEN?

Es fehlt weder an Ideen noch an Initiativen. Der Politik kann diesmal nicht der Vorwurf gemacht werden, nicht schnell und weitreichend genug ein Thema aufgegriffen zu haben. Offen bleibt, wie und welche der zahlreichen in der Politik diskutierten Regulierungsvorstellungen zu generativer KI von wem in die betriebliche Praxis übersetzt werden. Betriebsräte können sich auf zahlreiche inhaltliche Vorschläge berufen, wenn sie die Initiative übernehmen.

Vereinte Nationen: Zwischenbericht „Governing AI for Humanity“

Im November 2023 setzte UN-Generalsekretär António Guterres ein Beratungsgremium für den Umgang mit Risiken und Chancen in der Internationalen Governance von Künstlicher Intelligenz ein. Dieses „High-Level Advisory Body on Artificial Intelligence“ soll einen globalen, multidisziplinären und Multistakeholder-Dialog über die Regulierung von KI ermöglichen. Die Empfehlungen des Beratungsgremiums sollten die Vorbereitungen für den „Summit of the Future 2024“ der UN unterstützen und in die Verhandlungen über den vorgeschlagenen Globalen Digitalpakt (Global Digital Compact) einfließen.

Ende Dezember 2023 veröffentlichte das Beratungsgremium seinen Zwischenbericht „Governing AI for Humanity“. Der Bericht fordert eine engere Abstimmung zwischen internationalen Normen und der Art und Weise, wie KI entwickelt und eingeführt wird, und nennt sieben Prinzipien, die die Bildung neuer globaler KI-Governance-Institutionen leiten sollten:

- **Inklusivität:** Alle Bürger*innen, auch im Globalen Süden, sollten Zugang zu KI-Tools haben und diese sinnvoll nutzen.
- **Öffentliches Interesse:** Die Governance sollte über den Grundsatz der Schadensvermeidung hinausgehen und einen breiteren Rahmen für die Rechenschaftspflicht von Unternehmen, die KI entwickeln, einsetzen und kontrollieren, sowie für nachgeschaltete Nutzer definieren.
- **Zentralität der Data Governance:** KI-Governance kann nicht von der Governance von Daten und der Förderung von Data Commons getrennt werden.
- **Universell, vernetzt und Multistakeholder:** KI-Governance sollte der universellen Zustimmung von Ländern und Interessengruppen Vorrang einräumen. Sie sollte bestehende Institutionen durch einen vernetzten Ansatz nutzen.
- **Völkerrecht:** KI-Governance muss in der UN-Charta, den internationalen Menschenrechtsnormen und den Zielen für nachhaltige Entwicklung verankert werden.

Das Gremium macht Vorschläge dazu, was KI-Governance beinhalten sollt:

- die regelmäßige Bewertung des Stands der KI und ihrer Entwicklung;
- die Harmonisierung von Standards, Sicherheits- und Risikomanagementrahmen;
- die Förderung der internationalen Multistakeholder-Zusammenarbeit zur Stärkung des Globalen Südens,
- die Überwachung von Risiken und die Koordinierung von Notfallmaßnahmen;
- die Entwicklung verbindlicher Rechenschaftsnormen.

G7-Verhaltenskodex für Organisationen, die fortschrittliche KI-Systeme entwickeln

Im Entwurf der Leitlinien für fortgeschrittene KI-Systeme fordern die Regierungschefs der G7-Staaten Tests und Maßnahmen zur Risikominimierung während des gesamten Lebenszyklus‘ der Systeme:

- die Rückverfolgbarkeit von Datensätzen, Prozessen und Entscheidungen;
- Mechanismen zur Meldung von Missbräuchen und Dokumentation von gemeldeten Vorfällen;
- eine öffentliche Berichterstattung über die Grenzen derartiger Systeme, über angemessene und unangemessene Einsatzzwecke;
- eine verständliche Gestaltung von Transparenzberichten;
- Wasserzeichen oder andere Techniken, die es den Nutzern ermöglichen, KI-generierte Inhalte zu identifizieren;
- die Offenlegung von KI-Governance- und Risikomanagement-Richtlinien einschließlich Datenschutzstandards und Maßnahmen zur Risikominimierung;
- robuste Sicherheitskontrollen;
- die Entwicklung von Tools, mit denen Nutzer*innen feststellen können, ob bestimmte Inhalte durch fortschrittliche KI-Systeme erstellt wurden;
- die Priorisierung eines verantwortungsvollen Umgangs mit vertrauenswürdiger und menschenzentrierter KI;
- die Implementierung geeigneter Dateneingabekontrollen und Audits.

Stellungnahme des Bundesbeauftragten für den Datenschutz

Der Bundesbeauftragte für den Datenschutz und die Informationsfreiheit nahm die Gelegenheit wahr, am 24. Mai 2023 zum Thema generative Künstliche Intelligenz vor dem Digitalausschuss des Deutschen Bundestages Stellung zu beziehen. Dabei stellt er folgende Forderungen:

- Einhaltung der datenschutzrechtlichen Grundlage zum Schutz personenbezogener Daten bei Entwicklung, Training und Einsatz von KI-Systemen gleich welcher Art;
- Prüfung der Datenquellen, die in das KI-Modell eingeflossen sind;
- Verfügbarkeit von Fact-Checking-Mechanismen;
- einen risikobasierten Regulierungsansatz, um auf dynamische Entwicklungen zu reagieren;
- eine Pflicht zur Kennzeichnung von primär KI-generierten Medien, nach dem Schema, wie Attribution urheberrechtlich geschützter Bilder und Inhalte bereits heute praktiziert wird;
- Aufklärung und die Schaffung von Bewusstsein für einen verantwortungsvollen Umgang mit derartigen Techniken;
- Schaffung von Verantwortlichkeiten für KI-bedingte Schäden;
- eine Transparenzpflicht zu den in die Trainingsdaten eingeflossenen Datenquellen.



Stellungnahme des Bundesamtes für Sicherheit in der Informationstechnik (BSI)

Das Bundesamt für Sicherheit in der Informationstechnik (BSI) verlangte in der Anhörung im Digitalausschuss:

- eine Risiko- und anwendungsbezogene Regulierung;
- einen Nachweis für die Einhaltung der Regulierung und die dazu erforderlichen Prüfkriterien, Prüfmethoden und Prüfwerkzeuge;
- Werkzeuge, die einen Hinweis darauf geben, ob ein Text maschinengeneriert ist;
- Wasserzeichen, anhand derer maschinengenerierte Ausgaben identifiziert werden können;
- die Prüfung des Wahrheitsgehalts von KI-Outputs durch mindestens zwei unabhängige Quellen;
- eine transparente Gestaltung von KI;
- die Formulierung von Mindestanforderungen beim Einsatz in den unterschiedlichen Risikoklassen;
- Warninformationen für Verbraucher*innen.

Regulierungsvorschläge der Wissenschaft

In derselben Anhörung des Bundestages nahm **Professor Dr. Philipp Hacker LL.M. (Yale)** Stellung. Seinen umfassenden Ausführungen sind auch Hinweise zu betrieblicher Regulierung von generativer KI zu entnehmen.

Er adressierte u.a. folgende Anliegen:

- Sicherstellung eines angemessenen Niveaus an Leistung, Interpretierbarkeit, Korrigierbarkeit und Sicherheit während des gesamten Lebenszyklus eines Modells;
- Testing mit unabhängigen Sachverständigen;
- Risikofolgenabschätzung für Grundrechte, Umwelt, Demokratie und Rechtsstaatlichkeit.

Professor Dr. Natali Helberger zeigt in der Anhörung die Grenzen aller Transparenzbemühungen auf und ordnete diese nur als ersten Schritt ein, um Nutzer in die Lage zu versetzen, fundierte Entscheidungen zu treffen. Sie führt aus: „Echtes Empowerment verlangt, das den Endnutzern die Möglichkeit gegeben wird, Einspruch zu erheben, (z.B., wenn der Output schädlich oder irreführend ist), sich abzumelden (Opt-out), Feedback zu geben oder detaillierte Erklärungen zu verlangen, zum Beispiel um Vertrauenswürdigkeit, Genauigkeit oder die Folgen beurteilen zu können.“

Der Hamburgische Beauftragte für Datenschutz und Informationsfreiheit veröffentlichte am 13. November 2023 eine „Checkliste zum Einsatz LLM-basierter Chatbots.“ Darin finden sich konkrete Vorschläge für die Gestaltung der Bedingungen und die Regulierung von generativer KI:

- Formulieren Sie klare und dokumentierte interne Weisungen, ob beziehungsweise unter welchen Voraussetzungen welche Tools infrage kommen. Konkrete Beispiele der zugelassenen und der untersagten Einsatzszenarien helfen bei der Verdeutlichung.
- Beziehen Sie immer Ihre*n interne*n Datenschutzbeauftragte*n ein, wenn Sie interne Weisungen erstellen oder einen Anwendungsfall erstmals umsetzen. Je nach Anwendungsfall sollten Sie in dem Zuge eine Datenschutzfolgenabschätzung erstellen. Gegebenenfalls kann es auch sinnvoll sein, Betriebs- und Personalräte mit ins Boot zu holen.
- Wenn sich der Anbieter eines Chatbots in den Geschäftsbedingungen eine Verwendung für eigene Zwecke einräumen lassen, dürfen keine personenbezogenen Daten an die KI übermittelt werden. Das betrifft jegliche Informationen, die Rückschlüsse auf Kund*innen, Geschäftspartner*innen oder sonstige Dritte enthalten, und ebenso Daten der eigenen Beschäftigten.

- Achten Sie darauf, dass Ergebnisse der KI-Anwendung möglichst keine personenbezogenen Daten enthalten. Auch wenn der Eingabebefehl keine Person nennt, kann die KI unter Umständen vorherige Eingaben oder Informationen aus dem Internet einbeziehen.
- Stellen Sie berufliche Chatbot-Accounts zur Verfügung. Beschäftigte sollten nicht eigenständig und unter Verwendung privater Daten ein Konto erstellen.
- Nutzen Sie die Option, die Verwendung Ihrer Daten zu Trainingszwecken abzulehnen.
- Kontrollieren Sie, ob die Ergebnisse richtig sind.
- Überprüfen Sie die Ergebnisse auf Diskriminierung.
- Entscheidungen mit Rechtswirkung sollten grundsätzlich nur von Menschen getroffen werden.
- Sensibilisieren Sie die Beschäftigten durch Schulungen, Leitfäden und Gespräche dahingehend, ob und wie sie KI-Tools nutzen dürfen.

Stellungnahme des Deutschen Gewerkschaftsbundes

In einer ersten Einschätzung nahm am 11. Juli 2023 auch der Deutsche Gewerkschaftsbund zur Position des Europäischen Parlaments zum KI-Verordnungsentwurf Stellung. Der Gewerkschaftsdachverband greift viele der bereits oben genannten Forderungen auf und fordert darüber hinaus:

- eine umfassende technische Dokumentation in verständlicher Sprache;
- die Einrichtung eines Qualitätsmanagementsystems, das die Einhaltung der Anforderungen gewährleistet und dokumentiert;
- eine Kennzeichnungspflicht für KI generierte Inhalte.

Gesamtverband der Versicherer

Der Gesamtverband der Versicherer betont in einem Positionspapier, „eine Folgenabschätzung wäre an dieser Stelle ein wichtiges Instrument gewesen, um die Auswirkungen der künftigen Regulierung verlässlich beurteilen zu können.“ Der Bankenverband warnt unterdessen vor den Grenzen der Einsatzmöglichkeiten etwa von ChatGPT: „Im direkten Kontakt mit dem Kunden dürften die Einsatzmöglichkeiten von ChatGPT oder vergleichbaren Textrobotern hingegen auf kurze Sicht noch begrenzt sein. Denn in der unmittelbaren Kundenkommunikation muss sichergestellt sein, dass Aussagen oder gar Empfehlungen sachlich korrekt, intern nachvollziehbar und ethisch unangreifbar sind.“

„ChatGPT: Samsung verbietet Mitarbeitern die Nutzung generative KI“, lautete ein Pressebericht Anfang des Jahres 2023. Samsung hatte entdeckt, dass Programmcodes von Entwickler*innen der Firma auf ChatGPT hochgeladen wurden. Das Unternehmen hat daraufhin seinen Beschäftigten nach einem Bericht von Bloomberg (Paywall) offiziell untersagt, derartige Plattform zu verwenden. Etwas differenzierter gehen einige deutsche Unternehmen mit der Verwendung von ChatGPT um. Das Handelsblatt berichtete davon, dass der Drogeriemarkt dm ein firmeneigenes ChatGPT-System einsetzt, das auf einer Lösung von Microsoft beruht. Die Anwendung zielt offensichtlich darauf, die Vorteile generativer KI zu nutzen und dabei beispielsweise den Abfluss von Unternehmensdaten oder die Erzeugung von illegalen Inhalten zu verhindern. Währenddessen berichtet das Handelsblatt auch, dass die Deutsche Bank die Webseite von ChatGPT für ihre Beschäftigten gesperrt hat.



Betriebliche Gestaltungsbeispiele

Provinzialversicherung

Die Provinzialversicherung dürfte die erste Sozialpartei sein, die mit ihrem Betriebsrat einen geschützten, experimentellen Betrieb von ChatGPT unter dem Namen „Provinzial GPT Chat“ vereinbart hat. Dabei werden die Inhalte als computergeneriert gekennzeichnet und Verantwortlichkeiten für den Output definiert. Die Verantwortlichen sollen die Inhalte auf Integrität, Genauigkeit und Zuverlässigkeit prüfen. Die Eingabe personenbezogener Daten ist untersagt und Microsoft wurde die Aufgabe gegeben, zu verhindern, dass Anfragen funktionieren, die den Bereichen Hetze, sexuelle Inhalte, Selbstverletzung und Gewalt zugeordnet werden können. Währenddessen hat Microsoft keine Möglichkeit, die Eingaben aus der Provinzialversicherung für ein Re-Training der eigenen Modelle zu nutzen. Die Anfragen bleiben im geschützten Bereich der Provinzial und werden spätestens nach 90 Tagen gelöscht. Eine Meldestelle für nicht regelgerechte Inhalte und Ungenauigkeiten des Systems wird geschaffen und auch der Betriebsrat wird sich an den Modelltests beteiligen.

Deutsche Telekom

Einen ersten Handlungsrahmen zur Nutzung von ChatGPT gibt es auch bei der Deutschen Telekom. Dort wird die Nutzung betrieblicher Mailadressen gestattet, aber ausdrücklich daran erinnert, dass die internen Regelungen zum Schutz von Persönlichkeitsrechten einzuhalten sind. In dem ersten Papier heißt es: Ergebnisse, die von Künstlicher Intelligenz erzeugt werden, dürfen keinesfalls ungeprüft verwendet werden.

AUTOREN

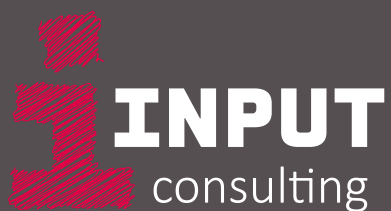
Matthias Spielkamp ist Mitgründer und Geschäftsführer von AlgorithmWatch, einer Menschenrechtsorganisation, die Auswirkungen von algorithmischen Systemen und KI auf Individuen und Gesellschaft kritisch beobachtet und begleitet. Er ist Mitglied im Beirat des deutschen Koordinators für Digitale Dienste (DSC), benannt durch den Deutschen Bundestag, und war von 2020 bis 2022 Mitglied der Global Partnership on Artificial Intelligence (GPAI).

Lothar Schröder war Sachverständiger der Enquete-Kommission Künstliche Intelligenz des Deutschen Bundestages, langjähriges Mitglied des ver.di Bundesvorstandes, wirkt als Vorsitzender des Datenschutzbeirats der Deutschen Telekom AG und berät Betriebsräte und Arbeitgeber zur Schaffung betrieblicher Ordnungsrahmen zum Thema Künstliche Intelligenz.

Markus Hoppe ist seit 2017 wissenschaftlicher Mitarbeiter bei INPUT Consulting und beschäftigte sich im Rahmen mehrerer Forschungs- und Gestaltungsprojekte mit den Auswirkungen des Einsatzes von KI-Systemen auf Beschäftigte und den daraus resultierenden Anforderungen an die betriebliche Mitbestimmung.

Karl-Heinz Brandl war bis Mai 2024 Gewerkschaftssekretär in der ver.di-Bundesverwaltung, u.a. als Bereichsleiter Innovation und Gute Arbeit sowie Projektleiter BeDax – Index Beschäftigten-datenschutz bei INPUT Consulting.

forschen | entwickeln | beraten



INPUT Consulting
Gemeinnützige Gesellschaft für Innovationstransfer,
Post und Telekommunikation mbH

Theodor-Heuss-Str. 2
70174 Stuttgart

Fon: +49 (0) 711 2 62 40 80

Mail: info@input-consulting.de

www.input-consulting.de

ISBN 978-3-910870-05-5