

Computer und Arbeit



cua-web.de

33. JAHRGANG
ISSN 1863-8511
D 11680

CuA | IT-MITBESTIMMUNG UND DATENSCHUTZ

5 | 2024

KÜNSTLICHE INTELLIGENZ

Microsoft Copilot im Fokus

MITBESTIMMUNG *Die technische Ausstattung des Betriebsrats*

ONLINE-TOOL *Datenschutz beim KI-Einsatz*

INFORMATIONSSICHERHEIT *Der Mensch als Lösung?*



KI und Ethik – im Dialog mit ChatGPT

KÜNSTLICHE INTELLIGENZ Der Einsatz fortschrittlicher KI-Systeme wie ChatGPT wirft viele Fragen auf. Wie kann ein fairer und ethischer Umgang mit der Technologie gelingen? Ein spannender Dialog mit ChatGPT selbst gibt Einblick in aktuelle Herausforderungen und Lösungsansätze.

VON ALEXANDER KLIER

Als in den 1960er Jahren erstmals die Diskussion um eine Ethik bezüglich künstlicher Intelligenz aufkam, ging es primär noch um die Anwendung von Algorithmen und programmierten Computern. Der amerikanische Informatiker Arthur Samuel war damals der Auffassung, dass diese Art von Maschinen die menschliche Intelligenz nicht übertreffen könne.¹ Allerdings, so fügte er hinzu, müsste dies neu bewertet werden, wenn Computer des Typs »neuronale Netze« verfügbar würden.

Mit den »Large Language Models« (LLM) wie ChatGPT stehen nun Systeme zur Verfügung, die sich durch ein so genanntes »deep learning« unter Verwendung ihrer neuronalen

Netze auszeichnen. Auch die Server, auf denen sie laufen, verfügen mittlerweile über genügend Ressourcen, um die Systeme in vielen Bereichen praktisch einzusetzen. Aber wie steht es denn nun mit den ethischen Implikationen eines solchen Einsatzes? Dazu interviewte ich eine KI, namentlich ChatGPT in der Version 4.²

Warum ethische Überlegungen?

Alexander Klier: »Warum sind ethische Überlegungen bezüglich einer KI notwendig?«

ChatGPT: »Ethische Überlegungen sind im Kontext von Künstlicher Intelligenz (KI) aus mehreren wichtigen Gründen erforderlich [...] KI-Systeme haben das Potenzial, viele

DARUM GEHT ES

1. KI-Anwendungen entwickeln sich derzeit sehr schnell.
2. Sie beeinflussen viele Lebensbereiche, auch die Arbeitswelt.
3. Wie können faire und ethische Ergebnisse sichergestellt werden?

¹ Samuel, Arthur (1960): Some Moral and Technical Consequences of Automation – A Refutation, Science Nr. 132, 9/1960, 741 ff.

² Sowohl die Fragen an wie auch die Antworten von ChatGPT sind in diesem Text unverändert wiedergegeben, mitunter allerdings gekürzt.

DEEP LEARNING

Deep Learning bedeutet auf Deutsch »tiefes Lernen«. Die »Tiefe« bezieht sich dabei auf die Schichten bzw. Ebenen, in der künstliche Neuronen hintereinandergeschaltet sind. Es handelt sich also um eine technische Bezeichnung, die zunächst nichts mit dem menschlichen Lernen zu tun hat.

Aspekte des täglichen Lebens zu beeinflussen, von der Arbeitswelt bis hin zur Freizeitgestaltung. Ohne ethische Leitlinien könnten KI-Anwendungen unbeabsichtigt negative Auswirkungen auf Individuen oder Gruppen haben [...] Zusammengefasst sind ethische Überlegungen bei der KI nicht nur nötig, um unmittelbare Herausforderungen anzugehen, sondern auch, um langfristig eine nachhaltige und positive Integration dieser Technologie in die Gesellschaft zu gewährleisten.«

Soweit eine sehr allgemeine und zutreffende Begründung dafür, warum man sich auf der ethischen Ebene diesem Thema nähern kann und auch nähern soll. Da ChatGPT in der Antwort betont, dass eine »Vertrauensbildung in KI-Anwendungen« wichtig sei, habe ich an dieser Stelle nach.

Alexander Klier: »Kannst du spezifizieren, wie Ethik und Vertrauen gewährleistet werden können?«

ChatGPT: »Die von KI-Systemen getroffenen Entscheidungen sollten nachvollziehbar und erklärbar sein. Erklärbarkeit (Explainability) sorgt dafür, dass Nutzer verstehen können, wie eine KI zu ihren Schlussfolgerungen kommt [...] Dies kann erreicht werden, indem die Modelle selbst so einfach wie möglich gestaltet werden oder indem zusätzliche Erklärungswerkzeuge bereitgestellt werden, die die Entscheidungsprozesse des KI-Systems auf eine verständliche Art und Weise interpretieren.«

Auch hier gibt ChatGPT den aktuellen Stand der Diskussionen korrekt wieder. Insbesondere die Erklärbarkeit ist eine wichtige und vertrauensfördernde Maßnahme. Es gibt mittlerweile eine eigene Disziplin, die verschiedene Methoden entwickelt, die die Ergebnisse von künstlicher Intelligenz und Algorithmen erklärbar machen sollen. Genannt wird sie »Explainable Artificial Intelligence« (XAI).

Fairness und Bias

Alexander Klier: »Welche ethischen Überlegungen legst du selbst zu Grunde?«

ChatGPT: »Als KI-basierte Assistenz kann ich selbst keine ethischen Überlegungen anstellen oder Entscheidungen treffen, da ich kein Bewusstsein oder Moral habe. Die »ethischen Überlegungen«, die ich erwähne, basieren auf den ethischen Grundsätzen und Richtlinien, die von Menschen für die Entwicklung und den Einsatz von KI-Systemen festgelegt wur-

den. Diese Prinzipien sind in mein Design und meine Programmierung integriert und dienen dazu, meine Interaktionen mit den Nutzern zu leiten.«

Ein für mich besonders interessanter Punkt sind die Fragen des immer wieder nachzuweisenden Bias. Unter einem Bias versteht man in den Sozialwissenschaften eine verzerrende Eigenschaft kognitiver Vorgänge. Bei allen Menschen besteht ein Bias, er ist psychologisch gut belegt. Speziell hierzu hatte ich folgende zwei Fragen gestellt.

Alexander Klier: »Wie wurde Fairness in dein Design integriert?«

ChatGPT: »Ohne ethische Richtlinien läuft KI Gefahr, bestehende soziale Ungleichheiten zu verstärken, indem sie Verzerrungen und Vorurteile aus den Daten, auf denen sie trainiert wird, repliziert [...] Die Daten, mit denen KI-Systeme trainiert werden, sollten auf Verzerrungen (Biases) geprüft werden. Ebenso sollten Algorithmen entwickelt werden, die aktiv Ungleichheit vermeiden und Gerechtigkeit fördern [...] Die Vielfalt menschlicher Kulturen und Werte bedeutet, dass KI so gestaltet sein sollte, dass sie universell akzeptabel und anwendbar ist, ohne dabei bestimmte Gruppen zu benachteiligen.«

Alexander Klier: »Kannst du verstehen, warum du nichts zu deinen eigenen Fairness-Richtlinien sagen kannst?«

ChatGPT: »Als KI-Assistent besitze ich kein Bewusstsein, Verständnis oder Selbstwahrnehmung im menschlichen Sinne. Ich bin nicht in der Lage, Konzepte wie Fairness zu begreifen oder über meine eigenen Richtlinien zu reflektieren. Ich kann keine unabhängigen Gedanken formulieren oder ein Bewusstsein für mein Design oder die implementierten Richtlinien und deren Auswirkungen entwickeln.«

Die gegebenen Antworten von ChatGPT treffen sehr gut den Problembereich. Interessant ist aus meiner Sicht auch, dass dieses Thema in den unterschiedlichen Kontexten von ChatGPT aufgegriffen wird, also nicht nur auf explizite Nachfrage erläutert wird.

ChatGPT zum Datenschutz

Alexander Klier: »Warum ist der Datenschutz für das Betreiben einer KI wichtig?«

ChatGPT: »Die Einhaltung von Datenschutzstandards wie der EU-Datenschutz-Grundverordnung (DSGVO) ist unerlässlich.

HINWEIS

Die Responsivität der menschlichen Sprache kann insofern ein Fallstrick sein, als wir in die Antworten gerne hineinlesen, was wir »hören« möchten. Das spiegelt sich auch bereits in der Formulierung der Fragen wider. Der Bias gilt am Ende also auch für die Fragen, die die Anwender:innen eingeben – und wie sie die Antworten interpretieren und aufnehmen.

Persönliche Daten sollen geschützt und nur mit Zustimmung der Betroffenen verwendet werden. Datenschutzmaßnahmen sollten in die Architektur der KI integriert sein (Privacy by Design) [...] Nutzer müssen darauf vertrauen können, dass ihre Daten sicher sind und verantwortungsvoll verwendet werden. Datenschutzrichtlinien und deren Einhaltung bauen dieses Vertrauen auf und ermöglichen es Nutzern, KI-Technologien zu akzeptieren und zu nutzen [...] Die ethische Nutzung von Daten ist ein grundlegendes Prinzip im Umgang mit KI. Datenschutz hilft dabei, ethische Standards einzuhalten und das Risiko von Diskriminierung, Voreingenommenheit und anderen ethischen Problembereichen zu minimieren.«

Obwohl ich diese Frage nur einmal explizit gestellt habe, ist auch das Thema Datenschutz permanent in den Antworten von ChatGPT vorgekommen.

Wie transparent sind die Ergebnisse?

Alexander Klier: »Wie kann ich nachvollziehen, aufgrund welcher Kriterien du zu diesen Punkten gekommen bist?«

ChatGPT: »Die von mir bereitgestellten Informationen und Empfehlungen basieren nicht auf dem selbstständigen Lernen oder der Entscheidungsfindung von künstlicher Intelligenz, sondern auf programmierten Algorithmen und bereits vorhandenen Datensätzen. Meine Antworten sind das Ergebnis der Verarbeitung von vorher festgelegten Informationen und Regeln, die von Entwicklern erstellt wurden und kontinuierlich aus großen Mengen verfügbarer Daten im Internet ergänzt werden, einschließlich wissenschaftlicher Publikationen, Fachartikel, etablierter Best Practices und allgemeiner Richtlinien zum Thema KI [...]. Solltest du spezifischeres Feedback oder eine tiefere Erklärung für einen der genannten Punkte wünschen, kann ich weitere Details liefern oder auf externe Quellen verweisen, wo die verwendeten Daten und Erkenntnisse im Detail nachgelesen werden können. Die Transparenz von KI-Systemen wie mir ist jedoch begrenzt durch die Komplexität der Algorithmen und den Umfang der verarbeiteten Daten, was ein vollständiges Verständnis aller zugrunde liegenden Prozesse für den Endnutzer oft schwierig macht.«

Quellen sollten beim Einsatz von ChatGPT immer geprüft werden. Hat man Ahnung von dem Gebiet, zu dem die künstliche Intelligenz

befragt wird, ist es zudem einfacher falsche Informationen als solche zu erkennen und herauszufiltern.

Empfehlungen für Betriebsräte

Alexander Klier: »Was würdest du Betriebsräten bezüglich des Einsatzes von KI empfehlen?«

ChatGPT: »Betriebsräte sollten sich für die Bereitstellung von Schulungs- und Bildungsressourcen für Mitarbeiter einsetzen, um sie auf den Umgang mit KI-Technologien vorzubereiten und das Verständnis für die neuen Arbeitswerkzeuge zu vertiefen [...] Es sollte deutlich gemacht werden, wie und warum KI im Unternehmen eingesetzt wird und welche Veränderungen das für die Belegschaft bedeutet [...] Das Gremium sollte analysieren, wie KI die Arbeitsbedingungen und -prozesse beeinflusst und sicherstellen, dass diese Veränderungen nicht zu einer Verschlechterung derselben führen [...] KI-Systeme sollten fair gestaltet sein und dürfen keine Mitarbeitergruppen benachteiligen. Der Betriebsrat sollte auf die Einhaltung dieser Prinzipien achten [...].«

Verantwortung für ethische Standards?

Verantwortlich für die Sicherstellung ethischer Standards sind, hier hat ChatGPT absolut recht, die »Entwickler, Unternehmen und Organisationen, die mich betreiben und warten. Sie setzen diese Prinzipien bei der Auswahl der Trainingsdaten, der Entwicklung der Algorithmen und der Überwachung meiner Leistung um.« Das bedeutet allerdings nicht, dass Nutzer:innen jegliche Verantwortung abgeben sollen. Der Einsatz von KI und auch die von ihr gelieferten Ergebnisse sind stets zu prüfen und kritisch zu hinterfragen.

Betriebsräte spielen laut ChatGPT »eine wichtige Rolle dabei, sicherzustellen, dass KI-Technologien auf verantwortungsvolle und faire Weise im Unternehmen implementiert werden« und »im Einklang mit den Interessen der Belegschaft stehen«. Dieser Aussage kann und will ich nichts hinzufügen, weshalb ich das »Interview« an dieser Stelle beende. ◀



Dr. Alexander Klier, Co-Founder und Enabler New Work & Change, Pforzheim und München.
www.alexander-klier.net

Guter Rat für Personalräte



Deppisch / Jung / Schleitzer

Tipps für neu- und wiedergewählte Personalratsmitglieder

Rechtliches Wissen und soziale Kompetenz
5., aktualisierte Auflage
2022. 254 Seiten, kartoniert
€ 24,90
ISBN 978-3-7663-7218-5

bund-shop.de/7218



service@bund-shop.de
Info-Telefon: 069/95 20 53-0