

06.08.2024 | Risikomanagement | Schwerpunkt | Online-Artikel

## Ethisches Handeln lässt sich nicht auf KI abwälzen

verfasst VON: Michaela Paefgen-Laß

7:30 Min. Lesedauer

### Jetzt neu: KI-gestützte Suche!

Markieren Sie Textabschnitte, um KI-gestützt weitere passende Inhalte zu finden.

powered by

Aktivieren Sie unsere intelligente Suche, um passende Fachinhalte oder Patente zu finden.

KI-gestützte Suche

Patentsuche

Aus

**Eine Künstliche Intelligenz ist eine Künstliche Intelligenz und damit nicht mehr als Technik und Statistik. Von ihr ethisches Handeln zu erwarten, lenkt ab von den wirklichen Verantwortungsträgern.**





Chat-GPT hat die Möglichkeiten Generativer KI in der breiten Öffentlichkeit erfahrbar gemacht. Damit ist auch die Diskussion um Ethik und Moral in der Anwendung von KI zu neuem Leben erwacht. © Atchariya / stock.adobe.com

**Hype, Horror oder Helfer: Auf Large Language Models (LLMs) basierende generative Künstliche Intelligenz (Generative KI) wie Chat-GPT polarisiert und scheint privaten Akteuren im Alltag so unvermittelt aufgetaucht, wie die sprichwörtlichen biblischen Plagen. Das schürt bei ihnen einen diffuses Unbehagen und befeuert dystopische Fantasien. Andere erliegen euphorisch den Verheißungen der KI und setzen übersteigerte Erwartungen in ihre autonomen Fähigkeiten. In den Führungsetagen großer Unternehmen und Arbeitsräumen von Start-ups läuft unterdessen die Auseinandersetzung mit den Chancen und Risiken der Generativer KI.**

## **Generative KI ist gekommen, um zu bleiben**

Die Zukunftstechnologie ist in deutschen Unternehmen angekommen und erfüllt dort ihren Zweck, indem sie Prozesse verbessert, Innovation vorantreibt, Entscheidungsfindung erleichtert und Aufgaben wie die Text- und Bildgestaltung automatisiert. Rund 96 Prozent der deutschen Unternehmen gehen einer aktuellen Erhebung des Bundesverbands Digitale Wirtschaft (BVDW) e.V. davon aus, dass sie mit Generativer KI ihre Geschäftsergebnisse verbessern können.

Die Versprechen der Generativen KI gehen in zwei Richtungen: Das Erweitern der menschlichen Fähigkeiten durch das Lösen komplexer Probleme und die Entlastung des Menschen, durch die Übernahme verschiedener – häufig repetitiver – Tätigkeiten. Aber wie ist es um die Technologiefolgenabschätzung bestellt? Auf dem Beipackzettel der Generativen KI stehen Gefahren und Risiken, die das Vertrauen von internen wie externen Interessengruppen aufs Spiel setzen. Wer übernimmt die ethische Verantwortung für die bekannten, die vorhersehbaren und die unerwarteten Nebenfolgen?

## **Generative KI birgt für Digitalexperten ethische Risiken**

Kognitive Technologien, wie die Generative KI, werden dem Deloitte

Technology Trust Ethics Survey 2023 zufolge von Branchenspezialisten und Digitalexperten weltweit als zweischneidiges Schwert gehandelt. So sind sich 39 Prozent der 1.700 befragten Digitalexperten darin einig, dass kognitive Technologien das größte Potenzial für den sozialen und gesamtgesellschaftlichen Nutzen haben. Gleichzeitig finden 57 Prozent, dass dort auch die größten Gefahren für ernsthafte digital-ethische Verstöße lauern.

In 65 Prozent der befragten Unternehmen wird Generative KI bereits intern eingesetzt und 31 Prozent haben damit begonnen, Generative KI für den externen Gebrauch zu nutzen. Weitere 74 Prozent sind in der Erprobungsphase. Da Generative KI noch nicht lange auf dem Markt ist, gilt der verantwortungsvolle Umgang mit ihr als größte Herausforderung für Anwender und Nutzer. Die digital-ethischen Bedenken der Befragten erweisen sich daher als eine Mixtur ethischer, technischer und rechtlicher Aspekte. Was ein Indiz für Komplexität der Technologie und die relative Unerfahrenheit im Umgang mit ihr ist.

### **Digital-ethische Bedenken, befürchtete Konflikte und Verstöße, beim Einsatz Generativer KI (Deloitte, 2023)**

- Privatsphäre: 22 Prozent
- Transparenz: 14 Prozent
- Manipulation: zwölf Prozent
- Geistiges Eigentum/Urheberrecht: zwölf Prozent
- Datenherkunft: zwölf Prozent
- Halluzinationen: neun Prozent
- Täuschung: acht Prozent
- Arbeitsplatzverlagerung: sieben Prozent
- Aktualität: drei Prozent

### **Worauf KI-Anwender achten, um Nutzer zu schützen**

Bei der Einführung von KI-Lösungen in Unternehmensprozesse gilt es auf einen verlässlichen und rechtssicheren Einsatz hinzuarbeiten. KI-Nutzer haben in der Regel wenig Wissen über die Funktionsweise einer KI-Lösung, weshalb die Anwendung für sie automatisch mit unbekanntem Risiken verbunden ist. Voraussetzung für den nachhaltigen Vertrauensaufbau ist also eine ethisch-orientierte Handlungsweise der KI-Anbieter.

Das Forschungsprojekt "Trust KI" am Institut für Internet-Sicherheit – if(is) der Westfälischen Hochschule hat es sich zur Aufgabe gemacht, als Mittler im Vertrauensaufbau zwischen Herstellerunternehmen und Anwenderunternehmen zu fungieren. Ziel ist, die Akzeptanz von KI-Anwendungen in der Breite zu verbessern.

Welche Kriterien erfüllt sein müssen, um ein Vertrauensverhältnis zwischen Führungskräften und der KI-Lösung sowie ihren Anbietern aufzubauen, hat das Institut mit seiner Anwender-Studie untersucht. Daran beteiligten sich 263 Führungskräfte deutscher Anwenderunternehmen, von denen sich 111 Teilnehmende "solide Kenntnisse" und 99 "Grundkenntnisse" im Bereich der KI bescheinigten

Vertrauenswürdigkeitsaspekt	Kernergebnisse der Trust KI-Studie
<b>Zutrauen</b>	<ul style="list-style-type: none"> <li>• für 50,8 Prozent ist "sehr wichtig", dass der KI-Anbieter hat seinen Hauptsitz in der EU/Deutschland hat.</li> </ul>
<b>Zuverlässigkeit</b>	<ul style="list-style-type: none"> <li>• 65 Prozent erwarten, dass der KI-Hersteller einen Prozess zur Qualitätssicherung in der Entwicklung definiert hat.</li> <li>• 75,4 Prozent erwarten, dass der KI-Hersteller einen Störfall unmittelbar kommuniziert</li> </ul>
<b>Integrität</b>	<ul style="list-style-type: none"> <li>• 62,2 Prozent wollen über die Grundsätze der Geschäftsbeziehungen informiert werden.</li> <li>• 61,3 Prozent haben Informationsbedarf zur sozialen, ökologischen und ökonomischen Nachhaltigkeit.</li> <li>• 50 Prozent erwarten die Einhaltung der ethischen Grundprinzipien Fairness und Solidarität</li> <li>• 72,6 Prozent wollen wissen, ob und wie sichergestellt wird, dass bei der Nutzung der KI-Lösung ein Anwender die Kontrolle über die Abgabe seiner Daten behält.</li> <li>• 68 Prozent wollen wissen, was der KI-Hersteller unternimmt, um physische, psychische oder finanzielle Schädigungen von Anwendern abzuwenden.</li> <li>• für 48,3 Prozent ist relevant, ob ein Ethik-Gremium im Unternehmen etabliert wurde und welche Verantwortung/Befugnisse es hat.</li> <li>• 54,9 Prozent ist es wichtig, ob und wie der KI-Hersteller eine Folgenabschätzung auf Anwender/Gesellschaft unternimmt.</li> </ul>

- 42,4 Prozent wollen wissen, ob und warum der KI-Hersteller auf bestimmte Funktionalitäten zum Wohle des Kunden verzichtet.

## **Die "Integrität" einer Generativen KI ist Sache von wem?**

Die Vertrauensdimension "Integrität" und die damit verbundene Verantwortungsübernahme für das Nicht-Schädigungsprinzip war den von Trust KI befragten Anwendern besonders wichtig. Akteure in der Debatte um Ethik und Moral im Umgang mit komplexer künstlicher Intelligenz sind die Wissenschaft, Wirtschaft und die Öffentlichkeit. Hier wird ausgetragen und geregelt, wer verantwortlich ist für Konflikte und Verstöße, die etwa durch algorithmische Diskriminierung entstehen.

Künstliche Intelligenz sei ein "inhärent konservatives Instrument", sagt Ethikprofessorin Judith Simon im Gespräch mit NDR-Journalist Jürgen Deppe. Dystopische Untergangsszenarien vernebelten den Status quo der Technologie, die aus alten Daten lerne und diese Muster dann in Prognosen und Empfehlungen fortschreibe. Simon ist Mitglied im Deutschen Ethikrat und Professorin für die Ethik der Informationstechnologie an der Uni Hamburg. Der Ethikrat untersucht die Auswirkungen digitaler Technologien auf das Verhältnis zwischen Mensch und Maschine und bewertet diese entlang ethischer Richtlinien.

## **Die Welt der Generativen KI ist "von gestern"**

Generative KI-Systeme, wie sie von Industrie, Verwaltung, Medizin sowie privaten Nutzern eingesetzt werden, sind Input-Output-Systeme, die durch Datentrainings neue Muster in "alten" echten oder synthetisch generierten Daten entdecken. Nachlässig eingespeiste oder überholte Daten verursachen eine "altmodische" Weltsicht der KI-Systeme und bergen das Risiko Verzerrungen zu reproduzieren.

In seiner Stellungnahme "Mensch und Maschine - Herausforderungen durch Künstliche Intelligenz" vom März 2023 verweist der Ethikrat daher auf die Datenqualität als entscheidenden Faktor für die Leistungsfähigkeit und Zuverlässigkeit von datengetriebenen Anwendungen:

Es geht also darum, dass Daten nicht nur prinzipiell von hoher oder niedriger Qualität sind, sondern auch für eine jeweilige Frage oder Aufgabenstellung mehr oder weniger angemessen sein können. Werden solche Fragen der Passung nicht rechtzeitig und hinreichend berücksichtigt, sind Verzerrungen oder irreführende Analysen möglich. (Deutscher Ethikrat, 2023/95)

## **Noch ist KI keine superintelligente AGI**

Bei Künstlicher Intelligenz handelt es sich um Systeme, die "die in einem eng definierten Kontext ebenso eng definierte Aufgaben und Probleme lösen" schreiben die Springer-Autoren Reinhard Kreissl und Roger von Laufenberg (Seite 256). Risiken und Gefahren drohten demnach nicht von der technologischen Leistungsfähigkeit der Systeme, sondern ihrer unkontrollierten Verbreitung. Künstliche Intelligenz hat kein eigenes Bewusstsein und kann sich keine eigenen Ziele stecken. Die überraschenden Leistungen von LLMs wie Chat GPT befeuern allerdings Spekulationen über den wirklichen Stand in der Entwicklung von superintelligenter, selbst gesteckte und übergeordnete Ziele verfolgender *artificial general intelligence* (AGI).

Wird eine menschenähnliche AGI schon bald in der Lage sein, Empathie und moralisches Urteilsvermögen zu entwickeln? Oder sollte Forschung, die sich zum Ziel setzt AGI zu erzeugen, schlichtweg verboten werden, wie Holger Hoos, Professor für Methodik der Künstlichen Intelligenz, Rheinisch-Westfälische Technische Hochschule Aachen (RWTH), in einem Press-Briefing durch Science Media Center Germany fordert.

Wir wollen KI, die uns hilft, unsere Schwächen zu erkennen und auszugleichen und die uns die Bereiche, in denen wir stark sind, im Wesentlichen weiterhin überlässt. Das ist meine Definition von menschenzentrierter KI und die ist kompatibel mit der der Europäischen Union. (Holger Hoos, RWTH Aachen)

## **Europäisches KI-Gesetz will Risiken von KI-Anwendungen regulieren**

Der Ethikrat fordert in seiner Stellungnahme, dass KI menschliche Entfaltung erweitern muss und nicht vermindern darf. Dass sie den Menschen, seine Intelligenz und Verantwortung nicht ersetzen darf. Ethische Standards für

eine vertrauenswürdige KI formuliert die KI-Verordnung (Artificial Intelligence Act) der Europäischen Union. Der AI Act verfolgt einen risikobasierten Ansatz.

Er verpflichtet alle Akteure entlang der Wertschöpfungskette von Künstlicher Intelligenz zur Verantwortungsübernahme. Die EU will mit dem im März 2024 verabschiedeten AI Act - dem nach eigenen Angaben weltweit ersten umfassenden KI-Gesetz - sicherstellen, dass KI Systeme "sicher, transparent, nachvollziehbar, nicht-diskriminierend und umweltfreundlich sind".

Im AI Act legt die EU Verpflichtungen für Anbieter und Nutzer fest. Sie verbietet Systeme mit unannehmbaren Risiken, definiert Hochrisiko-Systeme und formuliert Transparenzanforderungen, die auch für nicht als risikoreich eingestufte Technologien wie Chat-GPT gelten. Insgesamt werden KI-Anwendungen mit niedrigem Risiko und Start-ups weniger stark reguliert. Ziel ist, die Entwicklung und Akzeptanz von vertrauenswürdiger KI-Modelle zu fördern. Innerhalb von zwei Jahren muss das Gesetz in nationales Recht umgewandelt sein. Bereits ab Herbst kann das Verbot für KI-Systeme, die unannehmbare Risiken darstellen, gelten. Systeme mit hohem Risiko haben drei Jahre Zeit, die Anforderungen zu erfüllen.

 MyTopic Alert

Loggen Sie sich ein, um Ihre Alerts zu aktualisieren und Neue anzulegen.

Künstliche Intelligenz	Innovationsmanagement	
Datenmanagement	Transformation	Big Data
Datensicherheit	Industrie 4.0	Robotik
Automatisierung	Energieverteilung	

---

**Premium Partner**