



Geister kontrollieren, die wir riefen

RALF HOFSTETTER

Nutzung von **KI**: Ein schmaler Grat zwischen Wertschöpfung und Verantwortung.

Schon seit seiner breiten Verfügbarkeit im November 2022 hat das KI-basierte Dialogsystem Chat GPT von Open AI wiederholt Schlagzeilen gemacht – positive und negative. Zwar wurde die Sprachsoftware mit Schutzmechanismen versehen. Trotzdem liefert sie immer wieder Fehlinformationen, diskriminierende oder schablonenhafte Antworten. Nur weil maschinell lernende Anwendungen menschliche Fähigkeiten oft weit übertreffen, bedeutet das nicht, dass sie unfehlbar sind.

Das ist weder in deren begrenztem technologischen Potenzial noch in vermeintlich propagandistischen Absichten des Open-AI-Forschungsteams begründet. Vielmehr befindet sich Chat GPT derzeit in einer frühen Entwicklungsphase. In ihr treffen die enorm hohen Erwartungen der Nutzenden und die Grenzen der Technologie aufeinander – mit öffentlichkeitswirksamen Konsequenzen. Die Erkennungsrate und damit der Mehrwert von KI-Sprachsoftware hängt wesentlich von der Anzahl und Qualität der Trainingsdaten und -vorgänge ab.

Regulation hinkt hinterher

Fehler oder Qualitätsmängel in diesen Daten und unvermeidbare Generalisierungen können auf das Entscheidungsverhalten der Anwendung durchschlagen und sie fehlerhafte Schlussfolgerungen ziehen lassen. Mit anderen Worten: Maschinell lernende KI ist immer nur so gut wie die Daten, mit denen sie trainiert wurde und wird. Chat GPT greift auf einen Datenfundus ab 2011 zurück. Das sind also «nur» elf Jahre Wissen und Erfahrung in Datenform.

Es ist nachvollziehbar, dass der Ruf

nach klaren Regeln für die Umsetzung von KI immer lauter wird. Hier tut sich derzeit einiges. Mit Gesetzen wie der EU-Datenschutz-Grundverordnung (DSGVO) oder dem revidierten Schweizer Datenschutzgesetz (nDSG) haben gesetzgebende Instanzen bereits ins digitale Marktgeschehen eingegriffen. Weitere staatliche und private Initiativen sowie diverse Prüf-

Digialethische Lösungen sind weder schwarz noch weiss, sondern eine Balance.

siegel für KI-Anwendungen sind in Arbeit. Mit dem «Proposal for a regulation laying down harmonised rules on artificial intelligence» will die Europäische Union das Vertrauen in KI stärken. Und mit Algo.Rules können Entwicklerinnen und Entwickler auf eine praxisnahe Orientierungshilfe zurückgreifen.

Trotz (selbst-)regulierender Initiativen bleiben die Unternehmen in der Pflicht. Denn wer digitale Wertschöpfung betreibt, muss sicherstellen, dass sie gemeinwohlverträglich ist. Allerdings ziehen Wirtschaftlichkeit und Ethik häufig Spannungsfelder auf. Zum Beispiel kann die Achtung der Privatsphäre im Umgang mit Personendaten den betriebswirtschaftlichen Erfolg schmälern. Automatisierte Rekrutierungsprozesse können die Diversität der Belegschaft beeinträchtigen. Oder aber der Einsatz eines Chatbots kann zu Enttäuschungen führen, weil Kundinnen und Kunden lieber persönlich mit Menschen interagieren.

Digialethische Lösungen sind weder schwarz noch weiss, sondern eine fragile Balance zwischen betriebswirtschaftlicher Wertschöpfung und verantwortungsbewusstem Unternehmertum. Um diese Balance zu wahren, müssen die Unternehmen ihr Handeln an digital-ethischen Werten ausrichten und einen hauseigenen Wertekodex ausarbeiten. Nur so stellen sie ihre technologische Innovationskraft in den Dienst der Menschen und verpflichten sich dazu, Technologien und Daten sicher, souverän, nachvollziehbar und verantwortungsbewusst zu nutzen.

Eine digialethische Strategie geht aus der Vision des Unternehmens hervor. Sie basiert auf der Geschäftsstrategie und den unternehmerischen Kernwerten, allen voran der Integrität. Mit einer derart integrierten Strategie kann das Unternehmen sicherstellen, dass Menschen, Prozesse, Produkte und Technologien in jeder Zyklusphase innerhalb der Grenzen seiner Digialethik liegen. Ein modulares Rahmenwerk erleichtert den Strategieprozess (siehe Abbildung), da sich die Verantwortungsbereiche dem firmenspezifischen Kontext anpassen lassen.

Digitale Ethik prüfen

KI-Anwendungen sind als Produkte zwar nicht haftbar, unterliegen aber dennoch einer Rechenschaftspflicht. Deshalb ist es wichtig, dass unternehmenseigene Leitlinien zur Digialethik jederzeit einer Überprüfung standhalten.

Hier kommt die Prüfung ins Spiel. Ein Prüfteam kann zum einen die Governance des Unternehmens prüfen und sicherstellen, dass prozessuale und technologische Prozesse den definierten Ethikgrundsätzen entsprechen, richtig umgesetzt und nachvollziehbar offengelegt werden. Zum andern kann das Prüfteam mit einer entwicklungsbegleitenden Prüfung das ethische Strategiefit eines innovativen digitalen Produkts ge-

währleisten. Schliesslich können die Prüfenden algorithmisierte Entscheidungen überwachen und so deren Fehleranfälligkeit kontrollieren.

Die ethische Vertrauenswürdigkeit selbstlernender KI beginnt mit deren Entwicklung und dem datenbasierten Trai-

ning. Umso entscheidender ist es, dass ein Unternehmen davor digitale Ethikgrundsätze niederschreibt und seinem Innovationsprozess zugrunde legt. Mit der glaubhaften und von einem Prüfteam bestätigten Bindung an digitaletische Richtlinien hilft ein Unternehmen mit, die digitale Transformation verantwor-

tungsvoll und gemeinwohlverträglich zu gestalten. Es beugt auch Reputationschäden vor und trägt zudem wesentlich zum nachhaltigen Vertrauen seiner Anspruchsgruppen und der Gesellschaft in wertschöpfende Innovationen wie Chat GPT bei.



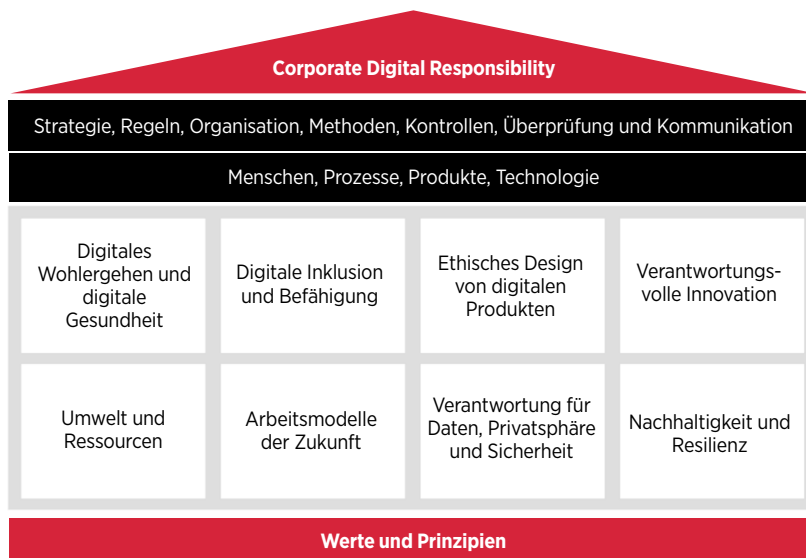
GETTY IMAGES/DIGITAL VISIONVECTORS

Werte und Prinzipien auch beim Einsatz von künstlicher Intelligenz: Das macht «Corporate Digital Responsibility» aus.

Modular aufgebaute Digialethik

Sie lässt sich gezielt in Unternehmensstrategie und Geschäftsprozesse integrieren

Ralf Hofstetter, Director Digital Assurance,
PwC Schweiz, Zürich.



QUELLE: PwC DEUTSCHLAND

