

Ethische Fragen im Umgang mit Künstlicher Intelligenz

Leitgedanken des Expertenrats Künstliche Intelligenz

Die Erkenntnis, dass es eines ethischen Rahmens für Künstliche Intelligenz (KI) bedarf, ist inzwischen ein elementarer Bestandteil der politischen Debatte rund um KI in Deutschland. Nicht nur Politik, Wissenschaft und Zivilgesellschaft, sondern auch zahlreiche Unternehmen nehmen ihre gesamtgesellschaftliche Verantwortung wahr und befassen sich schon seit einiger Zeit mit der Frage, auf welchen ethischen Grundlagen der Einsatz von KI basieren sollte.

Prinzipielle Einigkeit herrscht dabei hinsichtlich des übergeordneten Prinzips, dass die Steigerung des menschlichen Wohls Leitsatz für die Entwicklung und Anwendung von KI sein sollte. Dies hat die interdisziplinär zusammengesetzte „High-Level Expert Group on Artificial Intelligence“ (AI HLEG) der Europäischen Kommission in ihren im April 2019 vorgelegten ethischen Richtlinien für KI als „menschenzentrierten Ansatz von KI“ definiert.

Aufbauend auf diesem Grundsatz gilt es nun, ethisches Handeln in der betrieblichen Praxis bei der Entwicklung und Anwendung von KI zu verankern. Die Mitglieder des Expertenrats Künstliche Intelligenz¹ wollen hier ihrer Verantwortung, die sie in ihrem unternehmerischen Handeln gegenüber der Gesellschaft haben, gerecht werden und als Vorbild vorangehen. Der Expertenrat hat sich vor diesem Hintergrund gemäß seines Ziels, die unternehmerische Perspektive verstärkt in die aktuelle politische und gesellschaftliche Debatte rund um KI einzubringen, intensiv mit der Frage beschäftigt, wie die ethische Entwicklung und Anwendung von KI sichergestellt werden kann. Die daraus hervorgegangenen Leitgedanken hat er nachfolgend formuliert.

Die vorliegenden Schlussfolgerungen sind entstanden aus einer umfassenden Diskussion innerhalb des Expertenrats. Sie haben zum Zweck, grundlegende Prinzipien zu definieren, die als Orientierung für eine ethische Entwicklung und Anwendung von KI dienen. Die hier aufgeführten Leitgedanken erheben dabei nicht den Anspruch, erschöpfend oder abschließend zu sein. Ebenso bleiben die zum Teil weitreichenderen KI-Ethikkodizes, zu denen sich viele der im Expertenrat vertretenen Unternehmen bereits selbst verpflichtet haben, unbeschadet der hier genannten Grundsätze. Anspruch dieser Ausführungen ist vielmehr, ein von den Mitgliedern des Expertenrats geteiltes Fundament für den ethischen Umgang mit KI zu definieren.

1. KI sollte stets im Einklang mit dem westlichen Wertekanon entwickelt und angewendet werden.

Angesichts der unzähligen Einsatzmöglichkeiten von KI benötigen ethische Leitlinien für KI ein gemeinsames Fundament, das branchen- und produktübergreifend sowie über Ländergrenzen hinweg gültig ist. In der Vergangenheit diente das westliche Wertesystem, dessen zentrale Elemente auch im deutschen Grundgesetz und in der Europäischen Grundrechtecharta kodifiziert sind, den Unternehmerinnen und Unternehmern in Deutschland als ein solches ethisches Fundament. Aus Sicht des Expertenrats steht es außer Frage, dass diese Werte auch bei zunehmender Verbreitung von KI uneingeschränkte Gültigkeit besitzen und das unternehmerische Handeln in Deutschland weiterhin prägen müssen.

Für Unternehmerinnen und Unternehmer bedeutet dies in der Praxis, dass sie nicht nur Verantwortung für ihr eigenes Handeln tragen, sondern auch für dessen Auswirkungen auf ihr Unternehmen und die

¹ <https://www.microsoft-berlin.de/microsoft-expertenrat-fuer-kuenstliche-intelligenz.aspx>

Gesellschaft. Damit einher geht eine Verpflichtung zur Einhaltung ethischer Grundsätze. Zu diesen gehört u.a. eine Technologie – beispielsweise KI – so einzusetzen, dass sie der Gesellschaft durch die Schaffung eines wirtschaftlichen Mehrwerts sowie die Erhöhung des kollektiven und individuellen Wohlstands und Wohlbefindens dient.

2. Die Bewertung ethischer Fragen im Umgang mit KI sollte stets unter Betrachtung der Spezifika des betreffenden KI-Anwendungsfalls erfolgen.

Je nach Anwendungsbereich und Sektor variieren die spezifischen ethischen Herausforderungen von KI zum Teil beträchtlich. So wirft eine KI, die für die vorausschauende Wartung in der Industrie eingesetzt wird, andere ethische Fragen auf als diejenige, die beim automatisierten Fahren Anwendung findet.

Da viele der Anwendungsbereiche nur schwer vergleichbar sind, plädiert der Expertenrat deshalb für eine Einzelfallbetrachtung bei der ethischen Bewertung von KI-Anwendungen. Diese sollten die entwickelnden und anwendenden Unternehmen eigenständig und selbstverantwortlich vornehmen können. Grundlage und Orientierungsrahmen für diese Bewertung sollte das in Grundsatz 1 erwähnte Wertefundament sein.

3. Der Einsatz KI-basierter Anwendungen sollte in der direkten Interaktion mit Menschen stets kenntlich gemacht werden.

Die Vorstellung, in einer Welt zu leben, in der man nicht mehr klar zwischen Mensch und Maschine unterscheiden kann, verunsichert viele Menschen. Unwissentlich mit einer Computerstimme zu interagieren wirkt oftmals befremdlich. Die Tatsache, dass schon heute Bots mit Fake-Profilen zur Beeinflussung der öffentlichen Meinung in sozialen Medien eingesetzt werden, ist dagegen bereits eine manifeste Bedrohung für den gesellschaftlichen Diskurs sowie die freiheitlich-demokratische und wirtschaftliche Grundordnung unseres Landes.

In diesen und anderen Fällen den Einsatz von KI gegenüber den Nutzerinnen und Nutzern nicht kenntlich zu machen, kommt einem Täuschungsversuch gleich, der ethisch nicht vertretbar ist. Zugleich würde eine massenhafte Verbreitung solcher Täuschungen die Natur der Kommunikation und des gesellschaftlichen Zusammenlebens grundlegend verändern. Wer davon ausgeht, direkt mit einem menschlichen Gegenüber zu interagieren, nimmt den Austausch anders wahr und kommuniziert selbst auf eine andere Weise als mit einer Maschine. Wer fälschlicherweise glaubt, mit einer Maschine im Austausch zu stehen, läuft hingegen Gefahr, Grundregeln der zwischenmenschlichen Kommunikation zu missachten. Für die Verbraucher sollte es daher stets transparent gemacht werden, wenn sie mit einer KI interagieren.

4. Die einer KI zugrunde liegenden Entscheidungskriterien sollten im Rahmen der technischen Möglichkeiten nachvollziehbar gemacht werden.

Es muss sichergestellt werden, dass KI-basierte Systeme keine „Black Box“ sind, deren Kriterien für das autonome Treffen von Entscheidungen nicht einsehbar sind, wenn diese unmittelbare Auswirkungen auf das soziale Leben und die wirtschaftlichen Chancen von Individuen haben. Das aus gesellschaftlicher wie unternehmerischer Sicht so zentrale Vertrauen gegenüber KI-basierten Entscheidungen kann nur aufgebaut werden, wenn gewährleistet ist, dass diese genauso wie menschliche Entscheidungen geltendes

Recht respektieren und Entscheidungskriterien im Rahmen der technischen Möglichkeiten überprüfbar sind.

Zugleich liegt es in der Natur der einer KI zugrundeliegenden Algorithmen und KI-basierter Systeme, dass sie dynamisch und selbstlernend sind. Dies macht es technisch außerordentlich schwer umsetzbar, die einer spezifischen Entscheidung zugrundeliegenden Kriterien stets vollumfänglich nachvollziehen zu können. Dennoch gibt es technische Möglichkeiten, eine solche Nachvollziehbarkeit zumindest näherungsweise zu erreichen.

5. Um systematische Diskriminierung durch KI zu verhindern, sollte unter Einhaltung der bestehenden Datenschutzgesetze bei der KI-Entwicklung stets auf eine möglichst breite und diverse Datenbasis zurückgegriffen werden.

Die Nutzung einer möglichst breiten und diversen Datenbasis ist elementar, um das Aufgreifen bzw. Verstärken bereits bestehender systematischer Diskriminierungen durch KI-basierte Entscheidungsfindung zu verhindern. Im Bereich der Gesichtserkennungstechnologie wird beispielsweise häufig beobachtet, dass die Gesichter von Menschen mit heller Hautfarbe zuverlässiger erkannt werden als die von Menschen mit dunklerer Hautfarbe. Grund hierfür ist die Problematik, dass Datensätze häufig eine höhere Zahl an Bildern von Menschen mit heller Hautfarbe enthalten. Es ist daher essenziell, dass die zu verwendenden Datensätze die Charakteristika der relevanten Grundgesamtheit, aus der das Sample erhoben wurde, widerspiegeln und damit repräsentativ für diese sind. Dies darf selbstverständlich stets nur mit Daten vorgenommen werden, die im Rahmen der geltenden Datenschutzgesetze für diesen Zweck zur Verfügung stehen.

6. Um KI-Ethik langfristig in Unternehmen zu verankern, sollte sie als ein fester Bestandteil der Organisationskultur etabliert werden.

Die Einhaltung ethischer Standards auf der Anbieter- und Anwenderseite kann auf Dauer nur sichergestellt werden, wenn dies in der Unternehmens- und Organisationskultur fest verankert ist und von der Führungsebene selbst vorgelebt wird. Nur so kann sichergestellt werden, dass der Belegschaft der entsprechende ethische Rahmen gegenwärtig ist.

Wie dies in der Praxis umgesetzt werden kann, hängt stark von den existierenden Strukturen in den betreffenden Organisationen ab. Die Entscheidung darüber sollte demnach der jeweiligen Organisation obliegen.