

Guter oder böser Zauber?

Ethische Herausforderungen der KI

Text Dr. Elke Böckstiegel

Künstliche Intelligenz ist ein mächtiges Werkzeug – es kann erschaffen und es kann zerstören. Einerseits ist das Thema verbunden mit quasi-paradiesischen Aussichten auf Heilung, Reichtum und Müßiggang. Andererseits löst es Angst und Schrecken aus – vor Arbeitslosigkeit, totaler Entmündigung und Cyberwaffen. Stellen Sie sich KI als einen Zauberstab vor, mit dem man wie im Märchen wundersame Dinge bewirken kann: gute, aber auch böse.

Ich möchte hier **zwei ethische Problemfälle** der KI herausgreifen, mit denen wir als Gesellschaft und als Unternehmer umzugehen lernen müssen: den Missbrauch und die Entgleisung der KI aus dem von uns Gewollten.

Der **Missbrauch** ist der bekannte Fall. Macht ist notwendig als Gestaltungsmittel, aber gefährlich ohne Kontrolle. Als potentes Machtinstrument muss auch die KI Regeln unterworfen werden. Es braucht institutionalisierte Grenzen, etwa durch Gesetze und Verbote, und es braucht einen ethisch-moralischen Rahmen als Leitfaden für den Umgang mit dieser Macht. Nicht zufällig zog die Erfindung des Münzgeldes als mächtiges Instrument durch König Midas von Lydien im 7. Jahrhundert vor Christus unmittelbar die Ausbildung der Ethiken der großen Religionen und Kulturen in der sogenannten Achsenzeit nach sich, etwa bei den alten Griechen.



Insofern bringt die KI in diesem Punkt nichts prinzipiell Neues, sind die ethischen Regeln lediglich an die Möglichkeiten der KI anzupassen. Aber Vorsicht: Die damaligen Kulturen haben die Entfesselung der Kräfte durch das Geld erstaunlich weitsichtig erkannt und entsprechend reagiert. Heute ist es ungeheuer schwierig, die Auswirkungen der KI abzuschätzen – unser Vorstellungs- und Denkvermögen gerät an Grenzen. Zumal KI Dinge auf die Spitze treibt, etwa wenn sie von uns ethische Entscheidungen verlangt, die wir bisher dem Zufall überlassen konnten (z. B. Dilemmata in Unfallsituationen).

Der andere Fall – die **Entgleisung** – stellt sich als ethisches Problem neu. Denn: der Zauberstab KI ist nicht nur Werkzeug in unserer Hand, sondern denkt und handelt autonom. Wie können wir also verhindern, dass sich KI verselbständigt und Dinge tut, die wir nicht wollen und niemals beabsichtigt haben? Klingt verrückt? Ist es nicht. Aus zwei Gründen:

Grund 1: Schlechte Vorbilder. Bei einer KI zur Auswahl von Bewerbern in den USA stellte man verblüfft fest, dass sie nach kurzer Zeit diskriminierende und rassistische Entscheidungen traf. Warum? Man hatte sie zum Erlernen ihres Metiers mit den Daten der Vergangenheit, in diesem Fall mit den Einstellungsentscheidungen der letzten Jahre gefüttert. Die KI hatte daraus schnell gelernt, weiße jüngere Männer zu bevorzugen, jedoch Frauen, Ältere und Schwarze auszusortieren. Autsch!

Das ist ein großes Problem: Denn wir können nicht einfach sagen, füttern wir die KI doch mit guten Daten. Woher wollen wir wissen, welche Daten gut sind? Den Entwicklern der Rekrutierungs-KI war natürlich nicht klar, wie diskriminierend die Entscheidungen der Vergangenheit waren. Das wurde erst durch die KI wie durch ein Vergrößerungsglas sichtbar.

Noch schlimmer: Wir können gar nicht anders, als die KI mit Daten der Vergangenheit füttern. Aber wie gut ist die Vergangenheit? Wenn wir KI mit den Daten der letzten Jahrzehnte füttern, egal ob zu Finanzmärkten, Bildung, Klimaschutz oder Gesundheit, würde KI dann nicht die Irrwege der Vergangenheit fortsetzen, nur noch krasser, und so den jeweiligen Bereich schnell und final an die Wand fahren?

Grund 2: Unbeabsichtigte Folgerungen. Selbst wenn wir KI mit optimalen Daten fütterten, wüssten wir nicht sicher, was die KI daraus macht. Automaten ohne KI tun bei fehlerfreier Programmierung genau was sie sollen. Bei der KI können wir dagegen durch die Programmierung nicht exakt festlegen, wie die KI handelt – und aus welchem Grund.

Das Problem ist also nicht nur, a) wie wir die KI bewusst programmieren, sondern b) auch, mit welchen Daten wir sie füttern,

um zu lernen, und c) was die KI mit ihrer eigenen Logik daraus macht.

Fazit: Was folgt daraus? KI mag uns an vielen Stellen helfen, die Welt ein Stück besser zu machen. Aber sie ist selbst dann nicht automatisch ein guter Zauber, wenn wir sie mit bestem Wissen und Gewissen dafür programmieren. Denn in den benutzten Daten und Programmen stecken notwendigerweise auch alle unsere Unzulänglichkeiten, Fehler, Versäumnisse, blinden Flecken. Da wir die Wege der KI typischerweise nicht im Einzelnen nachvollziehen können, können wir das nicht allein durch mehr Sorgfalt und Kontrolle verhindern, sondern merken es erst an den Folgen. Dann kann es zu spät sein, weil uns das Tempo und die Mächtigkeit der KI überfordern.

Was tun? Wenn wir wollen, das KI uns mehr nutzt als schadet, müssen wir nicht nur träumen, sondern etwas tun! Sicherlich tun wir gut daran, KI immer mit einem Exit-Button zu bauen. Aber das reicht nicht. KI zwingt uns außerdem, massiv unsere Kompetenz zu verbessern, ständig Unvorhersehbares und Komplexität mitzudenken. Etwa sich lieber Schritt für Schritt stimmig entlang seines Anliegens vorzutasten statt blind nach Plan auf ein Ziel loszustürzen, das sich dann als fatal entpuppen mag. Solange wir das nicht der KI beibringen können, müssen wir es selbst meistern. Auch und gerade wir ehrbaren Kaufleute. Zum Glück kann man das lernen. Der Zauberstab KI kommt, und es ist höchste Zeit, uns auf den Weg vom überforderten Zauberlehrling zum souveränen Meisterzauberer zu begeben. Ob es uns gefällt oder nicht. Mit dem Turbo der KI im Rücken wird das existentiell. ■

Über die Autorin:

Dr. Elke Böckstiegel ist Mitglied im VBKI-Ausschuss Wirtschaft und Ethik. Sie ist geschäftsführende Gesellschafterin der Harmonia Logic GbR und Vorständin der Harmonia Academy AG.