

KÖNNEN ALGORITHMEN DISKRIMINIEREN?

| von WERNER ACHTERT, Geschäftsleitung Public Sector, msg

Künstliche Intelligenz (KI) hat das Potenzial, Einfluss auf das alltägliche Leben vieler Menschen zu nehmen. Deshalb findet aktuell abseits der technischen Diskussionen ein breiter gesellschaftlicher Diskurs über die Notwendigkeit einer stärkeren Transparenz beim Einsatz von KI-Systemen statt (siehe auch Seite 16 ff.).

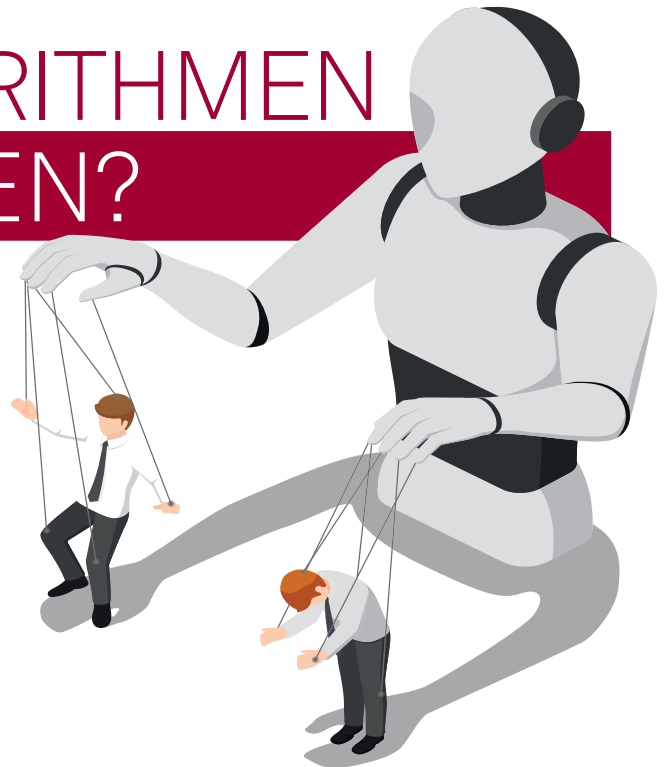
Ein zentrales Thema dabei ist der Vorwurf, KI-Systeme würden einzelne Personengruppen diskriminieren, indem zum Beispiel dunkelhäutige Menschen schlechter erkannt werden, Bewohner bestimmter Stadtteile bei Versandhäusern nur gegen Vorkasse bestellen können und Frauen bei Einstellungen schlechter bezahlt werden.

Dabei machen KI-Systeme eigentlich nur eines: Sie werten Daten der Vergangenheit aus und ziehen daraus – sehr vereinfacht ausgedrückt – durch Korrelationsanalyse Rückschlüsse zur Bearbeitung von weiteren Daten.

Was KI-Systeme nicht können, ist, Kausalzusammenhänge zu erkennen. Ein KI-System kann nicht „erklären“, warum zwei Mitarbeiterinnen oder Mitarbeiter bei vergleichbaren Tätigkeiten unterschiedliche Gehälter bekommen und schon gar nicht, ob das gerechtfertigt ist.

Wir dürfen bei der Diskussion über die Anwendung von KI nicht Ursache und Wirkung verwechseln. KI-Systeme schaffen keine gesellschaftlichen Probleme, sie können allerdings bestehende Ungerechtigkeiten, Benachteiligungen etc. offensichtlich machen. Der Einsatz von KI-Systemen birgt das Risiko, solche Verzerrungen zu verfestigen, aber auch die Chance, sie zu erkennen und zu beseitigen.

Gesellschaftliche Probleme können durch KI-Systeme nicht aus der Welt geschafft werden – sie können nur gesellschaftlich (also politisch) gelöst werden. In unserem Rechtssystem sind bereits Regeln zur Vermeidung von Diskriminierung vorhanden, wie beispielsweise das Allgemeine Gleichbehandlungsgesetz



(AGG). Diese können auf manuelle Entscheidungsprozesse genauso wie auf KI-Systeme angewendet werden.

Nicht das KI-System diskriminiert also, sondern die Auswahl der verwendeten Trainingsdaten, der Klassifikationen und der einer Bewertung zugrunde liegender Attribute (siehe auch Seite 22 ff.).

Was wir brauchen, sind mehr Sachverstand und gesellschaftliches Bewusstsein bei der Auswahl der Daten, mit denen wir KI-Systeme trainieren. Ein KI-System kann beispielsweise aus der Menge der eingegangenen Bewerbungen eine Auswahl für persönliche Einladungen treffen. Um Chancengleichheit zu gewährleisten, müssen in diesem Fall die Trainingsdaten ausgewogen in Bezug auf beispielsweise Alter und Geschlecht sein.

Wir müssen durch kontinuierliches Qualitätsmanagement, also zum Beispiel dem Hinterfragen der durch die KI getroffenen Entscheidungen, sicherstellen, dass die Trainingsdaten für den Einsatzzweck signifikant sind und keine ungewollten (oder gar gewollten) Verzerrungen enthalten. Und dies muss transparent gemacht werden, um bei den Betroffenen in der Öffentlichkeit Vertrauen zu schaffen.

Und dafür bedarf es – nach wie vor – menschlicher Intelligenz. ●