# „Manchmal ist es wichtiger zu wissen, was nicht in den Daten enthalten ist."

Immer neue Datenquellen und Künstliche Intelligenz treiben derzeit Gesellschaft, Wirtschaft und insbesondere die Marktforschung um. Alle müssen lernen, mit den neuen Möglichkeiten umzugehen und offen, aber auch kritisch zu bleiben. Teresa Kubacka über die Reproduzierbarkeit von Daten und die latente Überschätzung von KI.

Teresa Kubacka

***Teresa Kubacka*** ist Experimentalphysikerin und forscht im Bereich Künstlicher Intelligenz und maschinelles Lernen. Sie arbeitet an der Universität Zürich und hat sich einen Namen gemacht durch ihre Forschungen und ihre Kritik an den möglichen Gefahren von ChatGPT und anderen KI-Systemen. Ihre Hauptexpertise liegt im Bereich Data Science. Bevor sie sich als Data Scientist selbständig machte, arbeitete sie zwei Jahre lang als Data Scientist bei La Mobilière, einer führenden Schweizer Versicherungsgesellschaft. Dort hat sie an Projekten mit Computer Vision, NLP und Data Mining zur Verbesserung interner Prozesse gearbeitet.

**Teresa Kubacka, Sie haben einen Doktortitel in Physik. Wie kommen Sie von der Experimentalphysik zur Datenanalyse?**

Die Physik ist ein hervorragender Hintergrund für Datenwissenschaftler, denn beide Disziplinen haben die gleiche Kernkompetenz: die Modellierung der Vorgänge in der Welt mit Hilfe der Mathematik. Um ein erfolgreiches Experiment durchzuführen, muss ein Physiker die wichtigen und unwichtigen Faktoren verstehen, die die Messung beeinflussen können, und er muss wissen, wie er eine Datenerfassungsmethode entwickelt, um später das Signal vom Rauschen trennen zu können. Nehmen wir ein scheinbar einfaches Beispiel. Ich möchte die Temperatur des Wassers im Wasserkocher mit einem Temperatursensor messen. Einige Fragen, die ich mir stellen muss, sind: Wie zuverlässig und genau ist der Sensor? Wo soll ich den Sensor anbringen? Wie lange sollte eine einzelne Messung dauern? Wird mein Eingriff die Wassertemperatur verändern? Wie kann ich quantifizieren, wie vertrauenswürdig mein Ergebnis ist? Wenn Sie versuchen, die Daten selbst zu sammeln, lernen Sie sehr schnell, wie unübersichtlich und komplex die reale Welt ist und wie viel Sorgfalt Sie aufwenden müssen, um ein eindeutiges Ergebnis zu erhalten.

> *„Daten sind nicht Teil der natürlichen Welt."*

Man lernt auch, dass Daten nicht Teil der natürlichen Welt sind, sondern etwas, das von Menschen geschaffen wurde. Bleiben wir bei dem Beispiel: Wenn ich den Temperatursensor zu nahe an der Heizung platziere, erhalte ich keine zuverlässigen Daten über die durchschnittliche Temperatur im Wasserkocher. Wenn ich einen Sensor wähle, der nur zwei Werte misst: „heiß" und „kalt", kann dies die Daten für einige unvorhergesehene zukünftige Anwendungen unbrauchbar machen. Während also die Wassertemperatur selbst Teil der natürlichen Welt ist, ist meine Messung eine von Menschen gemachte Momentaufnahme der Welt, die stark von meinen Entscheidungen beeinflusst wird.

**Aber Physik und Marktforschung sind doch noch einmal deutlich anders.**

Natürlich gibt es erhebliche Unterschiede zwischen den beiden Disziplinen. Während ich in der Physik ein reproduzierbares Experiment entwerfen kann, bei dem ich vieles unter Kontrolle habe, haben wir es in der Datenwissen-

schaft mit unübersichtlichen, komplexen Daten zu tun, die eine multidimensionale, menschliche Realität widerspiegeln, in der wir leben. In diesem Sinne ähnelt die Datenwissenschaft viel mehr der Sozialwissenschaft. Auch die Ziele sind in der Regel sehr unterschiedlich: Während ich in der Physik ein robustes, vertrauenswürdiges Modell der Welt erstellen möchte, besteht das Ziel in der Datenwissenschaft in der Regel darin, ein gut funktionierendes, validiertes Blackbox-Tool zu haben, das eine bestimmte Aufgabe erfüllt.

> *„Hinter jedem Datenpunkt steht ein Mensch."*

**Wie sehen Sie neue Datenquellen oder Möglichkeiten zur Datengenerierung?**

Mein Traumszenario ist es, Daten in einer bekannten, kontrollierten Umgebung zu sammeln und dabei Tools zu verwenden, die gut verstanden werden. Im wirklichen Leben haben wir diesen Luxus nur selten. Die Sammlung und Aufbereitung von Daten ist eine äußerst ressourcenintensive Aufgabe, so dass wir oft über Daten verfügen, die „in freier Wildbahn" gesammelt wurden, oder wir verwenden Daten, die als Nebenprodukt eines anderen Prozesses gesammelt wurden, weiter. Dies bedeutet, dass die Daten alle Arten von Verzerrungen und unbeabsichtigten Korrelationen enthalten, die den endgültigen Einblick trüben können. Deshalb ist es notwendig, eine gehörige Portion Skepsis an den Tag zu legen und zu versuchen, all diese versteckten Faktoren zu verstehen, bevor man sich daran macht, Erkenntnisse zu gewinnen oder Produkte zu entwickeln.

Es ist auch äußerst wichtig, sich daran zu erinnern, dass hinter jedem Datenpunkt ein Mensch steht, der eine Entscheidung darüber getroffen hat, was wie gemessen werden muss und was nicht. Dieses Urteil wird von seinem Wissen, seinen Fähigkeiten und seiner Erfahrung, aber auch von seiner Sicht der Welt bestimmt. Zu wissen, was nicht in den Daten enthalten ist, ist manchmal sogar wichtiger als zu wissen, was enthalten ist.

**Was war Ihr erster Gedanke, als Sie ChatGPT getestet haben? Was sind Ihre ersten Fragen zu einem solchen System?**

Im Rahmen meiner Arbeit teste ich Tools zur Wissensentdeckung, die oft einige KI-Funktionen enthalten. Ich bin ChatGPT in ähnlicher Weise begegnet, mit Neugierde und Skepsis. Neugierde, weil KI-gestützte Tools, wenn sie richtig gemacht sind, erstaunliche Möglichkeiten bieten. Aber viele Tools – vor allem solche, die zu kommerziellen Zwecken entwickelt werden – versuchen, ihre Möglichkeiten zu überreizen. Da maschinelles Lernen bekanntermaßen

> *„Wie kann der KI-gestützte Missbrauch eingedämmt werden"*

gut darin ist, unbeabsichtigte Muster aufzuspüren, ist es umso wichtiger, diese Tools zu überprüfen und zu hinterfragen. Letztendlich geht es darum, durch den Umgang mit bekannten Risiken Vertrauen zu entwickeln. Für KI-Tools gibt es noch keine validierten Sicherheitsdatenblätter, so dass wir sie selbst hinterfragen müssen, um ein Gefühl dafür zu bekommen, ob wir uns auf sie verlassen können.

**Wie ist ChatGPT technisch einzuordnen?**

Die KI, die ChatGPT antreibt – ein Modell aus der Familie der Large Language Models (LLMs) – ist sehr schwierig, wenn nicht gar unmöglich, wahrheitsgetreu und sicher in der Anwendung zu machen. Das liegt daran, dass LLMs sich im Kern auf eine Verteilung von Wörtern aus Texten stützen, die sie bereits gesehen haben, und keine Vorstellung von einer „Tatsache" haben. Mit diesem Wissen war mein erster Schritt, zu überprüfen, wie wahrheitsgetreu ChatGPT in dem Bereich ist, in dem ich genug Wissen habe, um beurteilen zu können, wie zuverlässig das Ergebnis ist.

**ChatGPT und künftige andere Tools stehen für eine Demokratisierung der KI. Sollten KI-Systeme in der Gesellschaft allgemein verfügbar sein?**

Im Falle von ChatGPT und ähnlichen Produkten handelt es sich um eine paradoxe Demokratisierung. Einerseits werden in der Tat sehr leistungsfähige Werkzeuge für Menschen ohne technische Kenntnisse zugänglich, was eine ganze Welt neuer Möglichkeiten eröffnet. Andererseits werden diese KI-Systeme oft von internationalen Konzernen hinter verschlossenen Türen entwickelt, ohne externe Kontrolle und Transparenz. Dieses Spannungsverhältnis ist im Falle der LLMs besonders akut, die so enorme Mengen an Daten, geistigem Eigentum und Infrastruktur (Recheneinheiten, Strom usw.) erfordern, dass es kleineren Wettbewerbern und der akademischen Welt derzeit nicht möglich ist, sie zu replizieren. Es handelt sich also um eine Demokratisierung für uns als Verbraucher, aber nicht als Teilnehmer, die ein Mitspracherecht haben, wie diese Modelle erstellt werden, für welchen Zweck und zu welchen gesellschaftlichen und ökologischen Kosten.

**Gibt es Beispiele aus der Geschichte, wie wir hier den Weg bereiten können?**

Viele sagen, dass der Zugang zu KI in seinen wirtschaftlichen und gesellschaftlichen Auswirkungen mit dem Zugang zu Elektrizität vergleichbar ist. Wir könnten uns anschauen, wie wir gemeinsam dafür gesorgt haben, dass Elektrizität sicher und weithin verfügbar ist, und sehen, welche der daraus gezogenen Lehren wir nun auf KI anwenden können. Wir dürfen auch das große Ganze nicht vergessen, wo es viele offene Fragen gibt. Wie können wir den Wert, der sich aus der Produktion von Daten und der Nutzung von KI ergibt, aber auch die Gesamtkosten für die Entwicklung und den Betrieb dieser Systeme gerecht verteilen? Wie lässt sich ein Gleichgewicht zwischen Geschäftsgeheimnis und Gemeinwohl herstellen? Wie können die Menschen in die Lage versetzt werden, verschiedene KI-Modelle – nicht nur LLM – selbst zu entwickeln und sicher zu nutzen, und wie kann gleichzeitig der KI-gestützte Missbrauch eingedämmt werden? Es gibt keine einfachen Antworten, und wir müssen uns auf eine gründliche Diskussion einlassen, die bis zu den philosophischen Grundlagen dessen reicht, wie wir uns das Leben in einer Gesellschaft vorstellen.

*Die Fagen stellte Christian Thunig.*