

## Forum 2:

### „KI und unser Bild vom Menschen – Philosophische Perspektiven auf Künstliche Intelligenz“

Leitung: Dr. Martin Reilich, Cusanuswerk, Bonn

Referentin: Dr. Manuela Lenzen, Wissenschaftsjournalistin und Buchautorin, Bielefeld

Das Forum 2 der Jahresakademie zu künstlicher Intelligenz versuchte, die bereits bestehenden und in Zukunft durch KI noch relevanter werdenden Grenzen von Mensch und Maschine auszuleuchten. Die Wissenschaftsjournalistin *Dr. Manuela Lenzen* hat sich durch Veröffentlichungen zur Kognitionswissenschaft einen Namen gemacht, vor allem zur philosophischen Reflexion über Künstliche Intelligenz. Sie schreibt dazu Artikel und Monografien und bearbeitet auch als wissenschaftliche Mitarbeiterin am Zentrum für interdisziplinäre Forschung (ZiF) der Universität Bielefeld Fragen der Technikfolgenabschätzung.

In diesem Forum mit Teilnehmerinnen und Teilnehmern aus 17 Nationen lag der Schwerpunkt auf den sogenannten „letzten Bastionen“, d. h. auf denjenigen Eigenschaften des Menschen, die von KI nicht oder nur sehr unzureichend kopiert oder simuliert werden können: Kreativität, Moral, Bewusstsein, Autonomie.

In einem Rekurs auf die Geschichte und die optimistischen Anfänge zu Beginn der Entwicklung der künstlichen Intelligenz verwies Manuela Lenzen darauf, dass bei der berühmt gewordenen „Dartmouth Convention“ (1956, New Hampshire/USA) die Pioniere der einschlägigen Forschung noch davon ausgegangen waren, dass grundsätzlich „jeder Aspekt des Lernens oder jedes andere Merkmal der Intelligenz im Prinzip so präzise beschrieben werden kann, dass eine Maschine ihn simulieren kann“. Schon damals war das Ziel, Maschinen dazu zu bringen, „Probleme zu lösen, die heute dem Menschen vorbehalten sind, und sich selbst zu verbessern“.

Bevor das Forum in die Diskussion über die Unterschiede zwischen Mensch und Maschine einsteigen konnte, erklärte die Referentin einige technische Grundlagen, auf denen die moderne KI basiert und durch die sie in jüngster Zeit solch rasante Fortschritte machen konnte. Maschinelles Lernen basiert auf künstlichen neuronalen Netzen, die es maschinellen Lernmodellen (sog. „Foundation Models“) ermöglichen, aufgrund einer großen Menge von Daten derart trainiert zu werden, dass sie z. B. in Form von angepassten Chatbots eine Vielzahl von Aufgaben lösen können, die über eine Text-Eingabe (Chat) gestellt werden. Durch selbstüberwachtes Lernen oder halbüberwachtes Lernen entstehen viele Milliarden von Parametern, die dann als sogenannte „Transformer“ für angefragte Probleme eingesetzt werden können. Der weithin bekannte „Chat GPT“ („Generative Pre-trained Transformer“ der Firma Open AI) beispielsweise arbeitet mit 1,8 Billionen Parametern, der „Google Switch Transformer“ mit 1,6 Billionen.

Entscheidend ist aber, dass diese Lernmodelle sich genau in den Bereichen nicht selbst trainieren können, die Manuela Lenzen als „letzte Bastionen des Menschseins“ beschrieb: In Sachen Moral und Kreativität etwa sind die „Transformer“ darauf angewiesen, von Menschen trainiert zu werden, um durch dieses Training bessere Ergebnisse für die Nutzer generieren zu können.

An dieser Stelle richtet Manuela Lenzen den Blick auf den Global Süden, aus dem die meisten Teilnehmenden des Forums stammten. Gerade in Afrika werden zahlreiche junge Menschen als Arbeitskräfte für dieses Training eingesetzt. Es sind unzählige Stunden, Klicks und Einzelbeurteilungen nötig, um die riesige Masse von Parametern durch Menschen zu trainieren. In den Metropolen Afrikas finden sich hierfür genügend junge Arbeitssuchende, die dazu noch eine gute Schulbildung genossen haben. Wenn es um das Trainieren von moralischen Beurteilungen geht, ist diese Arbeit für die Betroffenen sehr belastend, weil sie täglich viele

Stunden lang gewaltförmigem und sexuellem Content ausgesetzt sind, denn der jüngste erstaunliche Boom des maschinellen Lernens (durch immer größer werdende Rechenkapazitäten, mehr verfügbare Daten und bessere Algorithmen) wird auf dem Rücken bedürftiger Menschen ausgetragen, die beim Trainieren der KI massiven psychischen Gefährdungen ausgesetzt sind.

Einer der wichtigsten „Take Home Messages“ des Forums war, wie sehr KI unser Bild vom Menschen verändert, weil sie gerade an den Rändern der maschinellen Möglichkeiten wie ein Spiegel fungieren kann, in den wir Menschen hineinsehen, um unser Wesen selbst zu erkennen. Das Bild von der Intelligenz entwickelt sich dabei weg vom reinen Akt des Verstandes und der Reflexion (symbolisiert durch Rodins Skulptur „der Denker“) hin zu einem kreativ-sozialen Geschehen des Konzipierens, aber auch des Ausprobierens und Interagierens (symbolisiert durch eine Gruppe Kinder, die gemeinsam einen Spielraum entdecken und erschließen).