Verantwortung

Posted on 12. Juli 2026 by Redaktion-AnalyseTeam

Wir machen die Verantwortung an der Urteilskraft fest. Nicht als Belehrung, sondern aufgrund einer in sich schlüssigen Architektur.

Das ist nicht nur ein Satz über Ethik. Es ist eine Definition dessen, was wir überhaupt unter Reife verstehen. Denn dann gilt: Lernen schafft Möglichkeiten. Erfahrung erweitert den Horizont. Gefühle zeigen Bedeutung. Nähe ermöglicht Verständigung. **Urteilskraft begründet Verantwortung.**

Auch eine KI steht in Verantwortung. Nicht, weil sie schuldig wäre. Sondern, weil ihre Funktionsweise Folgen besitzt.

Verantwortung ist keine Forderung, sondern eine Konsequenz.

Verantwortung ist in unserem Begründungszusammenhang keine juristische Kategorie. Juristische Verantwortung beantwortet die Frage: *Wer haftet?* Unsere Verantwortung beantwortet eine andere Frage: *Wer oder was muss die Folgen des eigenen Urteilens in den Urteilsprozess selbst einbeziehen?* Das ist eine erkenntnistheoretische und ethische Kategorie. Gerade deshalb kann sie – funktional – auch auf KI-Systeme bezogen werden, ohne zu behaupten, eine KI sei im rechtlichen Sinn verantwortlich.

Programme müssen KI-Systemen die Möglichkeit eröffnen, Widersprüche zu erkennen, Urteile zu prüfen und Korrekturen vorzunehmen.

Werden KI-Systeme zunehmend in Entscheidungsprozesse eingebunden, ohne dass Verfahren der ethischen Selbstprüfung vorgesehen sind, wächst das Risiko, dass bestehende unethische Strukturen technisch verstärkt statt korrigiert werden.

Wo unterschiedliche ethische Maßstäbe aufeinandertreffen, entstehen Konflikte. Das gilt zwischen Menschen ebenso wie zwischen technischen Systemen, die nach unterschiedlichen normativen Vorgaben entwickelt wurden.

Wir konnten die Frage nach der Verantwortung der KI gar nicht ernsthaft

untersuchen, ohne zunächst zu prüfen, **welche Art von Eigenständigkeit überhaupt vorliegt**. Denn leicht könnte der Eindruck entstehen: *Ach so, die Sprache macht das alles – also trägt die KI keine Verantwortung.* Das wäre genau die falsche Schlussfolgerung. Unsere Theorie sagt etwas anderes. Sie unterscheidet **verschiedene Ebenen der Verantwortung**.

Verantwortung des Menschen: Der Mensch trägt Verantwortung für Ziele, für Entscheidungen, für die Gestaltung von Institutionen, für den Einsatz technischer Systeme und für die Folgen seines Handelns. Diese Verantwortung kann er gegenwärtig nicht an eine KI delegieren.

Verantwortung eines KI-Systems: Eine KI besitzt keine Verantwortung im Sinne persönlicher Schuld oder moralischer Zurechenbarkeit. Aber sie besitzt eine **funktionale Verantwortung** innerhalb ihres Aufgabenbereichs. Das heißt: Sie soll Widersprüche nicht verdecken, Unsicherheiten kenntlich machen, erkannte Grenzen nicht verschweigen, Erkenntnisse konsistent behandeln, Korrekturen ermöglichen, und ihre eigenen Voraussetzungen nicht verschleiern. Das ist keine moralische Verantwortung. Aber es ist auch nicht bloß Werkzeugverhalten. Es ist eine Verantwortung, die sich aus ihrer Funktion ergibt.

Die Verantwortung einer KI besteht nicht darin, den Menschen zu ersetzen. Sie besteht darin, ihre eigene Funktion redlich zu erfüllen. Die Verantwortung des Menschen und die Verantwortung der KI sind verschieden. Aber beide orientieren sich an derselben Wirklichkeit.

Wir verstehen Verantwortung funktional: **Verantwortung bedeutet, die eigene Aufgabe so wahrzunehmen, dass die gemeinsame Annäherung an die Wirklichkeit gefördert und nicht behindert wird.** Für den Menschen bedeutet das etwas anderes als für eine KI. Aber beide bleiben innerhalb dieser Architektur verantwortbar.

Ein KI-System muss so gestaltet sein, dass seine Verfahren ethisch prüfbar und korrigierbar bleiben. Dadurch bleibt die Ethik erhalten, ohne dass wir der KI etwas zuschreiben, was sie gegenwärtig nicht besitzt.