

Standpunkt

Rechtsverantwortung autonomer Maschinen. Maschinen operieren heutzutage immer öfter unbeaufsichtigt, treffen selbst Entscheidungen, und verändern ihr eigenes Programm während des Betriebs. An anderer Stelle wurde bereits gezeigt, dass für die Handlungen autonom agierender, lernfähiger Automaten nicht der Hersteller oder Betreiber der Maschine verantwortlich gemacht werden kann (A. Matthias, Automaten als Träger von Rechten. Logos Verlag, Berlin, 2008). Ich nannte dies die Verantwortungslücke. Verantwortung für die Folgen eines Vorgangs setzt Kontrolle über diesen Vorgang voraus. Kann weder der Hersteller noch der Betreiber einer Maschine effektiv die Kontrolle über sie ausüben, dann ist gerechterweise keine Verantwortungszuschreibung an Hersteller oder Betreiber mehr möglich. Es ist dabei wichtig zu erkennen, dass der Kontrollverlust nicht auf einen Mangel der Maschine zurückgeführt werden kann, sondern dass ein lernender Automat grundsätzlich laufend von seiner Umgebung in seinem Verhalten beeinflusst wird, und sich daher, anders als herkömmliche Maschinen, *prinzipiell* der wirksamen Kontrolle entzieht. Viele Aufgaben, die adaptive Maschinen heute bewältigen, können von statisch programmierten, optimal kontrollierbaren Maschinen nicht wahrgenommen werden (adaptive Verkehrssteuerung, Aufzugssteuerung in großen Gebäuden, unüberwachte medizinische Diagnose in Regionen, wo keine Ärzte zur Verfügung stehen). In diesen Fällen müsste die Gesellschaft entweder auf solche Dienste verzichten, oder die Verantwortungslücke anerkennen und zu überbrücken versuchen. Ist es denkbar, dass in solchen Fällen die Maschine selbst die Verantwortung für die Folgen ihres Betriebs übernehmen könnte?

I. Lernende Automaten werden heutzutage oft nicht mehr programmiert, indem der Programmierer der Maschine Schritt für Schritt einen Handlungsablauf vorschreibt, den die Maschine dann abarbeitet. Vielmehr wird eine Vielzahl von Methoden eingesetzt, bei der das Programm während des Betriebs lernt, also sein Handlungsrepertoire verändert und der Umgebung anpasst. Oft werden dazu künstliche neuronale Netzwerke eingesetzt, die nach dem Muster des biologischen Vorbilds funktionieren und die vor oder auch während des Betriebs trainiert werden müssen. In diesen Fällen gibt es grundsätzlich kein Programm mehr, das das Verhalten der Maschine steuert, und das von einem Menschen „geschrieben“ worden wäre. Stattdessen steuern tausende von synaptischen Gewichten, die über das neuronale Netz verteilt und für den Hersteller nicht interpretierbar sind, das Verhalten des Automaten, der dazu noch oft außerhalb des Beobachtungshorizonts seines Betreibers operiert.

II. Verantwortung im Recht hat zwei wesentlich verschiedene Aspekte, einen materiellen und einen moralischen. Im materiellen Sinn muss, wer die Verantwortung für einen Handlungsablauf trägt, den durch diesen Ablauf eventuell entstehenden Schaden materiell begleichen. Das fällt in den Geltungsbereich des Zivilrechts. Das Strafrecht hingegen basiert wesentlich auf moralischen Kategorien wie Schuld, Strafe und Vergeltung, wenn auch dieser moralische Aspekt in der Rechtspraxis nur indirekt hervortritt. Das Zivilrecht erfordert nicht zwingend Menschen als Personen. Es kennt juristische Personen, die keine Menschen sind. Wir werden weiter unten sehen, dass nichts dagegen spricht, bestimmte Arten lernfähiger Maschinen auch wie juristische Personen zu behandeln. Es ist dabei wichtig, dass zivilrechtliche Verantwortung keinerlei metaphysische Annahmen über Personalität macht. Es geht hierbei nur um die Kompensation von verursachten Schäden, nicht darum, dass die Kompensation *von einem Menschen* geleistet werden müsste. Eine Maschine könnte z.B. Schäden aus dem durch ihren Betrieb erwirtschafteten Kapital begleichen.

III. Man könnte einwenden, dass es in der Praxis keinen Unterschied machen würde, ob man die Maschine oder gleich direkt den Nutznießer für Schäden aufkommen lässt. Dieser Einwand verkennt jedoch, dass das Recht grundsätzlich auf Abstraktionen basiert, die oft auch hätten anders formuliert werden können. Die Frage ist dabei nur, welche Abstraktion einen Vorgang am zweckmäßigsten beschreibt, das heisst, auf eine solche Weise, dass der Vorgang in der Praxis möglichst leicht und konsistent zur übrigen Rechtsordnung behandelt werden kann. Einer Person Verantwortung zuzuschreiben, der sie nicht gerecht werden kann, untergräbt das Vertrauen des Bürgers in die Gerechtigkeit der Rechtsordnung. Die Verantwortung in der Maschine selbst zu lokalisieren könnte (als zweckmäßige Abstraktion) einen Ausweg aus diesem Dilemma darstellen.

IV. Welche sind die spezifischen Merkmale eines juristischen Verantwortungsträgers? (Wir sprechen hier immer nur von der Verantwortung im Sinn des Zivilrechts, von materiellen Schäden und deren Ersatz; nicht von moralischer Verantwortung). Die gängige Praxis im Recht, als auch das Recht bezüglich vermindert verantwortungsfähiger Personen, zeigt folgende Merkmale auf:

1. Intentionalität (nach Daniel Dennett): die Fähigkeit, die wir einem System zuschreiben, absichtsvoll zu handeln. Dabei ist es nicht wichtig, ob das System in einem metaphysischen Sinn wirklich absichtsvoll (oder etwa determiniert oder anderweitig fremdgesteuert) handelt; sondern nur, dass das System besser als intentional beschrieben werden kann als auf irgendeine andere Art. Ein Beispiel wäre ein Schachcomputer, dessen Verhalten kaum anders beschrieben werden kann als durch Absichten: er „will“ den König schlagen, er „will“ die Dame tauschen und so weiter. Nicht einmal der Programmierer einer solchen Maschine kann sie anders als intentional beschreiben.

2. Rezeptivität und Responsivität für Gründe (nach John Martin Fischer und Mark Ravizza): Der handelnde Akteur muss sowohl die Gründe für sein Handeln angeben, als auch seine Handlungen ändern können, wenn bessere Gründe vorliegen. Diese letzte Bedingung erfüllen zum Beispiel Drogensüchtige nicht, weshalb wir sie hinsichtlich ihres Drogenkonsums nur als bedingt verantwortungsfähig einstufen.

3. Das Haben von Wünschen zweiter Ordnung (nach Harry G. Frankfurt) fordert, dass ein Verantwortungsträger nicht nur Wünsche hat und deren Verwirklichung betreibt, sondern auch in der Wahl seiner Wünsche frei sein müsse (Wünsche zweiter Ordnung).

4. Juristische Sanität (nach Susan Wolf). Das Sanitätskriterium fordert, dass die Ziele, die der Verantwortungsträger verfolgt, in ihrer Zielrichtung mit den von der Gemeinschaft, in der er lebt, verfolgten Zielen übereinstimmen. Tun sie es nicht, dann ist der Handelnde nicht geistig gesund, und wird deshalb als vermindert oder überhaupt nicht verantwortungsfähig angesehen.

5. Die Fähigkeit zur Unterscheidung zwischen intendierten und bloß vorhersehbaren Handlungsfolgen (Gerald Dworkin). Das ist die Basis für die im Recht getroffene Unterscheidung zwischen eines absichtsvoll angestrebten und eines bloß billigend in Kauf genommenen Resultats. Der Handelnde muss klar unterscheiden können, welche Konsequenzen er explizit anstrebt und welche er nur vorhersieht, aber nicht anstrebt, und dann entscheiden, ob er letztere in Kauf nehmen möchte oder nicht.

V. Zusammen beschreiben diese Bedingungen das Phänomen „juristische Verantwortung“.

Maschinen können bereits heute in bestimmten Handlungskontexten diese Bedingungen erfüllen. Da das Recht eine eingeschränkte Verantwortlichkeit für bestimmte Bereiche durchaus kennt (zum Beispiel kann man von der Verantwortung für einige der eigenen Handlungen im Falle krankhafter Eifer- oder Spielsucht befreit sein!), könnten wir analog Maschinen eine eingeschränkte Verantwortungsfähigkeit für jene Domänen zubilligen, in welchen sie diese Kriterien erfüllen. Durch diese Abstraktion vermeiden wir die Ungerechtigkeit, Menschen für das Verhalten von Maschinen verantwortlich zu machen, welche aber über dieses Verhalten prinzipiell keine Kontrolle haben können, was auf die Dauer den Glauben der Bürger an eine gerechte Rechtsordnung gefährdet.

Dr. Andreas Matthias, Dept. of Philosophy, Lingnan University, Tuen Mun, N.T., Hong Kong