

Moral für Maschinen?

Grundsätze der Maschinenethik



In vielen Anwendungsbereichen wie beim autonomen Fahren, im Krieg, aber auch in der Pflege stehen künstliche Systeme moralischen Problemen gegenüber. Die Maschinenethik ist eine aufstrebende Disziplin an der Schnittstelle von Philosophie, Informatik und Robotik, die sich mit der Frage beschäftigt, ob und wie künstliche Systeme auch moralische Entscheidungen treffen können. In diesem Beitrag wird die neue Disziplin vorgestellt und Argumente für und wider dieses Vorhaben diskutiert. Abschließend werden drei Grundsätze eingeführt, die als Leitplanken dienen und sicherstellen sollen, dass die technologische Entwicklung in eine positive Richtung erfolgt.

1. Was ist Maschinenethik?

Während *Artificial Intelligence* zum Ziel hat, die kognitiven Fähigkeiten von Menschen zu modellieren oder zu simulieren, geht es im Feld der *Artificial Morality* darum, künstliche Systeme mit der Fähigkeit zu moralischem Entscheiden und Handeln auszustatten. Die Idee ist also, Computer so zu programmieren, dass sie moralische Entscheidungen treffen können. Die Maschinenethik ist diejenige Disziplin, die sich mit der Möglichkeit von *Artificial Morality*, ihren theoretischen Grundlagen und ihrer ethischen Bewertung auseinandersetzt (Misselhorn 2018).

Lange Zeit stand die Maschinenethik zu Unrecht im Verdacht, bloß Science-Fiction zu sein. Das stimmt jedoch nicht. Denn bereits ein so einfaches Gerät wie ein Staubsaugerroboter steht vor moralischen Entscheidungen: Soll er einen Marienkäfer einfach einsaugen oder soll er ihn verscheuchen bzw. umfahren? Und was ist mit einer Spinne? Soll er sie töten oder ebenfalls verschonen? Ein solcher Roboter ist in einem minimalen Sinn autonom, weil er im Unterschied zu einem konventionellen Staubsauger nicht von einem Menschen geführt oder überwacht wird. Die Pointe eines solchen Roboters ist, dass er möglichst dann aktiv werden soll, wenn wir gerade nicht zu Hause sind.

Man kann auch in Zweifel ziehen, dass es sich um eine moralische Entscheidung handelt. Das ist jedoch der Fall, weil es darum geht, ob man Tiere zu Reinigungszwecken töten darf. Gewöhnliche Staubsaugerroboter können eine solche Entscheidung allerdings noch nicht treffen. Zumindest als Prototyp liegen jedoch Forschungsansätze vor, ein Ethikmodul für das populäre Modell Roomba zu entwickeln (Bendel 2017), welches das Leben von Insekten berücksichtigt (der Prototyp besitzt wahlweise einen ‚Kill-Button‘ für Spinnen).

 Pflegesysteme, Kriegerroboter und autonome Fahrzeuge stellen drei zentrale Anwendungsfelder der Maschinenethik dar

In komplexeren Einsatzbereichen autonomer Systeme stellen sich auch anspruchsvollere moralische Entscheidungen. Beispiele hierfür sind etwa Pflegesysteme, Kriegerroboter und autonome Fahrzeuge, die drei zentrale Anwendungsfelder der Maschinenethik darstellen (Misselhorn 2018). Alle drei Felder erfordern grundlegende moralische Entscheidungen, in denen es



Catrin Misselhorn

manchmal sogar um Leben und Tod von Menschen geht. Eine zentrale Frage der Maschinenethik ist, ob man Maschinen solche Entscheidungen überhaupt überlassen darf oder ob man es vielleicht sogar tun sollte.

Wir wollen zunächst Argumente zugunsten der Behauptung betrachten, dass wir eine Maschinenethik brauchen. Sodann werden Argumente dargestellt, die gegen die Maschinenethik geltend gemacht werden können. Diese Argumente beziehen sich auf der einen Seite auf die technische Machbarkeit und auf der anderen auf die moralische Wünschbarkeit von Maschinen-Moral. Wie sich erweisen wird, haben beide Seiten ihre Berechtigung. Deshalb kommt es darauf an, wie die Maschinenethik ausgestaltet wird. Ein Vorschlag dafür wird mit drei Grundsätzen der Maschinenethik vorgelegt, die den Einwänden Rechnung tragen und sicherstellen sollen, dass sich die Maschinenethik in eine gute Richtung entwickeln kann.