

# THEOLOGISCHE REVUE

117. Jahrgang

– Januar 2021 –

---

**Loh, Janina: Roboterethik.** Eine Einführung. – Berlin: Suhrkamp 2019. 241 S. (Suhrkamp-Taschenbuch Wissenschaft, 2277), brosch. € 18,00 ISBN: 978-3-518-29877-0

Nicht wenige Analysen der Coronavirus-Pandemie des Jahres 2020 halten u. a. fest, dass dieses globale Ereignis einen Schub in der Digitalisierung und Robotisierung aller unserer Lebensbereiche mit sich bringen wird. Schulen und Hochschulen sind in besonderer Weise davon betroffen. Aber natürlich bringt die Pandemie die Ersetzung von Menschen durch Maschinen nicht erst hervor, sondern diese Robotisierung ist längst in weite Bereiche unseres Alltags eingedrungen und hat in den letzten Jahren insbes. in den Feldern des autonomen Fahrens, der Pflegerobotik und der autonomen Waffensysteme zahlreiche ethische Studien stimuliert. Wie wir mit Robotern umgehen sollen, ist zu einer zentralen moralischen Frage des 21. Jh.s geworden.

Die Philosophin Janina Loh, die an der Univ. Wien seit Jahren zu diesem Thema forscht, hat nun eine der ersten monographischen Einführungen in die Roboterethik vorgelegt. L. versteht dabei Roboterethik als „eine Teilbereichsethik der Maschinenethik“ (10; 16), welche im deutschsprachigen Raum insbes. durch Oliver Bendel und Catrin Misselhorn bekannt geworden ist. Aber im Gegensatz z. B. zur einführenden, aber sachorientierten Darlegung in BARTNECK / LÜTGE / WAGNER: *Ethik in KI und Robotik* (München 2019) fokussiert sich L. auf zwei Bereiche: Die Frage des (moralischen) Status von Robotern (Abschnitt 2) und die Frage, inwiefern Roboter Verantwortungssubjekte und -objekte sein können (Abschnitt 3). Dabei stellt sie zahlreiche Ansätze aus der wissenschaftlichen Literatur vor und gegenüber – teilweise auch deutlich kritisierend. Die im engeren Sinne angewandt-ethischen Fragen, also ob z. B. Sexroboter erlaubt werden sollten und wenn ja, in welchem Rahmen, bearbeitet sie nicht. Dass der Einsatz von auf sog. künstlicher Intelligenz beruhenden Robotern (man denke an das „maschinelle Lernen“) auch große risikoethische Abwägungen notwendig macht, kommt kaum in den Blick. Die jeweils dargestellten Positionen zu den zwei Leitfragen werden auf Beispiele bereits existierender robotischer Technologie bezogen, u. a. auf den humanoiden Roboter „Kismet“, den Segensroboter „BlessU-2“ oder die Seehundimitation „Paro“. Ihren eigenen theoretischen Hintergrund bildet dabei eine „kritisch-posthumanistische“ Perspektive, die sich der „Inklusion“ auch vom Menschen geschaffener Artefakte in die moralische Welt verschreibt und damit versucht, den „Essentialismus“ des humanistischen und „vorrangig westlichen, weißen heterosexuellen, aristotelischen, männlichen und eben exklusiven“ (41) Menschenbildes aufzubrechen. – Dass „inklusive Theorien“ nur eine „Ergänzung“ (vgl. 197; 201f) zu herkömmlichen Ansätzen sein sollen, ist schwer einsehbar.

Abschließend plädiert L. für einen „offene[n] und kritische[n] Diskurs (211), für den sie offenkundig auch nichtmenschliche Wesen (212) als Diskurspartner mit aufzunehmen bereit ist. Insofern ist es wohl Angelegenheit des Diskurses selbst, über die Diskurspartner zu befinden – was vermutlich einen nicht abschließbaren Prozess bedeutet.

Sachlich stellen sich hier zwei ganz grundsätzliche Fragen: Wie soll man in einem inkludierenden Diskurs mit Diskurspartnern umgehen, die nicht inkludieren, sondern am überkommenen „Essentialismus“ festhalten? Werden diese dann exkludiert? Und: Kann man einem gewissen „Essentialismus“ für ein ethisches Überlegen überhaupt entkommen? Mit der Überwindung des Essentialismus geht konsequenterweise auch eine Abschaffung von Opfern verfehlten Handelns einher. Wenn Wissenschaft und Ontologie – wie für Donna Haraway und Karen Barad skizziert (vgl. 190; 192) – immer schon politisch sind, sind sie auch Spielball der Macht, und Ethik wird zu einem Instrument der Macht.

L. selbst interpretiert auch die beispielhaften robotischen Systeme in einem Sinn, dass diese noch ein umgrenztes „Sein“ haben, wenn sie z. B. vergleicht, welche Autor/inn/en sich für „Paro“ als moralisches Handlungssubjekt aussprechen würden und welche nicht (Tabelle S. 73). Aber ist die vorzugsweise bei demenzkranken Patienten eingesetzte Seehundimitation Paro überhaupt ein Gegenstand und nicht vielmehr eine Software? Eine Pflegeeinrichtung kann zehn physisch gleiche Paros vorhalten und sie immer wieder anders an die Patient/inn/en verteilen, solange sie nur auf die jeweilige Programmierung für Patientin A oder Patient B eingestellt werden. Wenn man die ganz konkrete Frage stellt, ob man einen Paro zerstören und vernichten darf, so wäre die grundlegende Überlegung in der Argumentationslinie dieses Buches zunächst, ob Paro ein „moralisches Handlungsobjekt“ ist. Auch für diese Frage stellt L. Positionen gegenüber (Tabelle S. 95). Aber was hat man zerstört, wenn man den Roboter zerstört hat? Die Software, die auf Patientin A zugeschnitten ist (vielleicht „bottom-up“ durch maschinelles Lernen erzeugt), bleibt in anderen Paros oder in der „Cloud“ erhalten. Es sind dann doch solche Probleme, zu denen man noch gerne Genaueres erfahren würde.

L. fordert auch zur eigenen ethischen Positionierung auf, aber welche Kriterien hat man als Sich-Positionierender denn zur Hand? Wenn die Kriterien selbst im Fluss des diskursiven Aushandlungsprozesses immer wieder neu ermittelt werden müssen, bleibt die Frage, ob nicht übermorgen vielleicht Personalität im Diskurs unter Robotern als moralischer Gesichtspunkt abgewählt werden könnte, sodass sie legitim die Menschen (wie in manchen Science-Fiction-Filmen) bekämpfen und ausrotten dürften. Die Ausweitung der Grenze bis zu artifiziellen Gegenständen wird letztlich um den Preis einer Auflösung der moralischen Position des Menschen erkaufte, was im Posthumanismus auch bewusst in Kauf genommen oder sogar angezielt ist. Aber die Ablehnung des Essentialismus betrifft nicht nur den Anthropomorphismus, sondern auch „Pathozentrismus“, „Biozentrismus“ oder „Physiozentrismus“ (110). Vielleicht kann sich eine „objektsexuelle“ Person geschlechtlich an ihr Handy binden, wie auf S. 82 dargestellt wird. Aber darf diese Person, falls ihr Handy gemeinsam mit einem Kind vom Schlauchboot fällt, das Handy retten und dafür das Kind ertrinken lassen? – Sind alle moralischen Normen und Werthierarchen nur historische Momentaufnahmen? L. sieht den „radikale[n] Sozialkonstruktivismus“ nicht als „einzige Alternative zu einem radikalen Objektivismus“ (124). Aber wäre es nicht spannend gewesen, hier konkretere Ortsbestimmung zu geben?

Der stark referierende und kommentierende Charakter der Darlegungen macht die Lektüre teilweise mühsam. Nicht alle Autor/inn/en verwenden die zentralen Begriffe wie „human agency“ oder „Verantwortung“ im gleichen Sinne, was Verwirrung stiften kann, was L. durchaus bewusst ist: In „Kritischen Zwischenbilanzen“ (121–125; 198–203) werden die Erträge der Debatte aus Sicht der Vf.in knapp geschildert, und tabellarische Übersichten fassen dann die in der Literatur diskutierten Positionen im Hinblick auf bestimmte Kriterien zusammen. Dennoch kommt es durch diese Vorgehensweise auch zu zahlreichen Wiederholungen, wenn sie z. B. sowohl auf S. 54 wie auf S. 144 feststellt, dass John Sullins den Ebenenunterschied von Verantwortung gegenüber Intentionalität und Autonomie übersehe, oder wenn sowohl auf S. 136 wie auf S. 146 mit Verweis auf Günter Wilhelms von den „behelfsmäßigen Begrifflichkeiten“ wie „Systemverantwortung“ die Rede ist, welche Verantwortung ohne (personales) Verantwortungssubjekt konzeptionell greifbar machen sollen. Um solche Schwächen in den Griff zu bekommen, hätte dem Buch vermutlich ein Lektorat durch eine Leserin oder einen Leser ganz gut getan, die oder der nicht so tief in den Debatten bewandert ist wie die Vf.in. Dass man aus der Darstellung sehr viel über die Diversität der roboterethischen (oder eher roboterontologischen) Debatten lernen kann, gehört zweifellos zu den Vorzügen des Buchs.

Über den Autor:

*Bernhard Koch*, Dr., PD, Stellvertretender Direktor des Instituts für Theologie und Frieden Hamburg (koch@ithf.de)