

Edward Ashford Lee

# **CO-EVOLUTION**

Die Symbiose von Mensch und Maschine

*aus dem Englischen übersetzt von  
Andreas Wirthensohn*

mandelbaum *verlag*

Gefördert durch

 **Bundesministerium**  
Klimaschutz, Umwelt,  
Energie, Mobilität,  
Innovation und Technologie

Titel der englischsprachigen Originalausgabe:  
*The Coevolution: The Entwined Futures of Humans and Machines*  
by Edward Ashford Lee  
Chapter 13 (*Pathologies*) and Chapter 14 (*Coevolution*)  
Copyright © 2020 MIT Press

mandelbaum.at • mandelbaum.de

ISBN 978-3-99136-027-8  
© mandelbaum verlag, wien • berlin 2023  
alle Rechte vorbehalten

Lektorat: ELVIRA M. GROSS  
Satz: KEVIN MITREGA, Schriftloesung  
Umschlag: MICHAEL BAICULESCU  
Druck: PRIMERATE, Budapest

# Inhalt

- 7 Vorwort
- 13 Erster Teil: Pathologien
- 51 Zweiter Teil: Co-Evolution
- 95 Anmerkungen



# Vorwort

Digitalisierung und Künstliche Intelligenz (KI) werden unser Leben grundsätzlich verändern: Welche Herausforderungen gehen damit für unsere Gesellschaft einher? Wie können wir diesen Prozess als Menschen und für die Menschen gestalten? Und was wird es in Zukunft überhaupt heißen, »Mensch zu sein«? Das sind fundamentale Fragen, denen sich seit 2022 ein gemeinsames Fellowship-Programm des Instituts für die Wissenschaften vom Menschen (IWM) und der Technischen Universität Wien widmet und in dessen Rahmen der vorliegende Essay von Edward A. Lee in deutscher Übersetzung beim Mandelbaum Verlag erscheint – mit der Absicht, weitere solcher Grundlagentexte folgen zu lassen.

Die Informatik, ihre Artefakte und Methoden, beeinflusst unsere Gesellschaft und damit unsere Welt, von der Ebene der Individuen bis hin zu den sich gegenwärtig beschleunigenden geopolitischen Machtspielen. Die Informationstechnologie (IT) dient heute als Betriebssystem unserer Gesellschaft. Aus ontologischer Sicht prägt die Informatik die Art und Weise, wie wir die Welt wahrnehmen und wie wir über sie denken. Dies zeigt sich auch in der rasch sich entwickelnden Künstlichen Intelligenz, der »letzten« Welle der fortschreitenden Digitalisierung.

Auch wenn momentan vielleicht nur zu erahnen ist, wie KI unser Leben konkret verändern wird, so ist doch gewiss: Die Einschnitte werden massiv sein und sich in vielen Bereichen bemerkbar machen. Schlagworte wie »Jobkiller«, »Wahrheitsvernichter«,

»Kreativitätseleminierer« und »Urheberrechtsverletzer« machen die Runde, um die Auswirkungen der KI zu skizzieren. Selbst wenn niemand seriös zu prognostizieren vermag, wie berechtigt solche Sorgen und wie weitreichend die möglichen Konsequenzen der KI sein könnten, sind mit diesen Schlagworten sicher einige der grundlegenden gesellschaftlichen Herausforderungen benannt.

Der Wandel der Arbeitswelt wird sich fortsetzen und die mit der Digitalisierung verbundene Vierte Industrielle Revolution sich weiter beschleunigen. Massiv datenbasierte Forschung könnte die Produktivität der Wissenschaft enorm steigern, bietet aber auch kaum noch zu übersehende Angriffsflächen für Kommerzialisierung und Manipulation. Bedroht sind überdies die politischen und gesellschaftlichen »Wahrheiten«, so dass insbesondere demokratische Prozesse anfällig erscheinen. KI erzeugt Texte, Gedichte, Bilder, Filme, Kompositionen, die einen großen Teil der sogenannten kreativen Tätigkeiten überflüssig machen könnten. Nicht zuletzt sind die komplexen juristischen Folgen des Einsatzes und der »Verantwortlichkeit« von KI ebenso wie ihres Rückgriffs auf Quellen und intellektuelle Leistungen, die von Menschen erzeugt wurden, nicht einmal ansatzweise zu überschauen. All das sind schon für sich genommen enorm anspruchsvolle Zukunftsfragen. Fundamentalere noch ist aber die Infragestellung des Menschseins: Was bedeutet es für das menschliche Selbstbild, dass Maschinen menschliche Leistungen kopieren, ersetzen und in vielen Bereichen überbieten können? Wie werden sich Menschen überhaupt noch als eigenständig von den Maschinen abgrenzen?

Edward A. Lee gibt darauf im vorliegenden Essay eine überraschende Antwort: gar nicht! Seine Vorstellung einer »Co-Evolution« – titelgebend für diesen Band – besagt, dass Mensch und Maschine gemeinsam einen Evolutionsprozess durchmachen werden: »Die daraus resultierenden Menschen werden sich womöglich stark von den heutigen Menschen unterscheiden.« Diese These mag provokant erscheinen, entspringt aber Lees Über-

zeugung, dass nur auf dieser Grundlage die zukünftige Technologieentwicklung noch beeinflusst werden kann. Alle bisherigen Versuche, die schädlichen Auswirkungen der Technologie einzudämmen, sind seiner Überzeugung nach wirkungslos geblieben, weil sie schlechte Entwicklungen als das Ergebnis des unmoralischen Handelns von Einzelnen sehen. Es gilt aber, so Lee, die evolutionäre Gesamtentwicklung in den Blick zu nehmen und den technologischen Prozess »anzustupsen«, sodass sich die Entwicklung in eine für den Menschen positive Richtung vollzieht. Die wechselseitige Interdependenz von Mensch und Maschine ist hier weder verklärte Utopie noch Schreckensszenario, sondern – in Teilen schon heute existierende – Lebensrealität.

Wir erleben eine Co-Evolution von Technologie und Mensch, die menschliche Interaktionen, gesellschaftliche Institutionen, die Wirtschaft und politische Strukturen grundlegend verändert – und auch die Wissenschaft ist davon natürlich nicht ausgenommen. Diese Entwicklung hat zwei Seiten: Neben ihrer scheinbaren Unausweichlichkeit aufgrund der Erweiterung von Möglichkeiten zum Funktionieren unserer Gesellschaft hat sie auch eine Reihe negativer Folgen wie etwa das fast automatische Erzeugen von Fake News und die damit zusammenhängenden Filterblasen, die zunehmende Überwachung und die damit einhergehenden Fragen des Datenschutzes, die kritische Veränderungen in der Arbeitswelt, die hohe Konzentration im Web mit anstehenden Antitrust-Aufgaben, die Abnahme der Souveränität von Einzelstaaten in ICT-Fragen, die drohende Einschränkung menschlicher Entscheidungsfindung durch automatisierte Systeme, KI-gesteuerte automatische Waffensysteme oder auch negative Umweltauswirkungen von IT-Systemen.

Die »doppelgesichtige« Rolle der IT, ihre unbestreitbaren enormen Errungenschaften und Potenziale und gleichzeitig die offensichtlichen kritischen Fragen, die miteinander verknüpft sind, war auch die Motivation für den ersten Wiener Workshop zum »Digitalen Humanismus« an der TU Wien im April

2019. Als intellektueller Ausgangspunkt diente Poppers Theorie von der Verantwortung der Wissenschaft, die dazu auffordert, Technologien nach menschlichen Werten und Bedürfnissen zu gestalten, anstatt zuzulassen, dass Technologien den Menschen formen. Der Workshop war auch von der Tradition des Wiener Kreises inspiriert, einem multidisziplinären Versuch des frühen 20. Jahrhunderts, über die revolutionären Auswirkungen der Wissenschaft auf unser Verständnis der Welt nachzudenken.

Seither geht es bei allen Diskussionen im Rahmen der Initiative zum *Digital Humanism* um die vielschichtigen Beziehungen zwischen Informatik und Gesellschaft. Es zeigt sich immer wieder, dass die Informatik zwar wichtig ist, aber allein nicht ausreicht, um umfassende Antworten auf die Herausforderungen zu geben; vielmehr ist ein breiter interdisziplinärer Ansatz gefragt. Die Teilnehmer der Initiative sind davon überzeugt, dass es möglich ist, diese Entwicklungen zu beeinflussen, ja dass es in unserer Verantwortung liegt, dies zu tun.

Auf dem Workshop im Jahr 2019 wurde auch das Wiener Manifest für digitalen Humanismus verabschiedet, es liegt in der Zwischenzeit in acht Sprachen vor und hat Unterzeichner aus fast fünfzig Ländern (cf. [caiml.dbai.tuwien.ac.at/dighum/dighum-manifesto](http://caiml.dbai.tuwien.ac.at/dighum/dighum-manifesto)). Das Manifest ist ein Aufruf zum kollektiven Handeln, um Unterstützung über nationale Grenzen und Kontinente hinweg zu mobilisieren und so eine menschlichere und nachhaltigere Zukunft zu schaffen. Der Begriff »Digitaler Humanismus« ist bewusst gewählt, um auf die Konzepte des Humanismus und der Aufklärung zu verweisen, wonach der Mensch für seine Handlungen und Überzeugungen verantwortlich ist und im Mittelpunkt steht. Dabei liegt die Betonung auf der Bedeutung des rationalen und kritischen Denkens als Verweis auf den Wiener Kreis und seinen logischen Empirismus.

Menschen haben die Freiheit, das Recht und die Verantwortung, von ihrem Denken Gebrauch zu machen, und sind selbst die Autoren ihres Lebens. Persönliche Autonomie und

Entscheidungsfreiheit sind die Voraussetzungen für eine offene, demokratische und ökologisch nachhaltige Gesellschaft. Der technologische Fortschritt ist weder gottgegeben, noch folgt er einem technischen oder ökonomischen Determinismus. Wir als Individuen und als Gesellschaft sollten und müssen Entscheidungen unter Berücksichtigung demokratischer, humanistischer und ökologischer Überlegungen treffen. Der Digitale Humanismus definiert sich demgemäß als ein Ansatz, der das komplexe Zusammenspiel von Technologie und Mensch beschreibt, analysiert und vor allem beeinflusst, um so eine bessere Gesellschaft und ein besseres Leben zu gestalten.

Einer der Erstunterzeichner und ein wichtiger Mitinitiator der Digital-Humanism-Initiative war Edward A. Lee. Wir sind sehr glücklich, nicht nur dass er der erste Visiting Fellow im Rahmen unseres neuen Programms war, sondern wir danken ihm auch für die bereitwillige Kooperation bei dieser deutschsprachigen Publikation seines bemerkenswerten Essays, von dem wir glauben, dass er entscheidende Impulse für zukünftige Debatten bereithält.

Unser besonderer Dank für die Unterstützung des Fellowship-Programms gilt dem Österreichischen Bundesministerium für Klimaschutz, Umwelt, Energie, Mobilität, Innovation und Technologie (BMK), namentlich Michael Wiesmüller, Leiter der Abteilung Schlüsseltechnologien für industrielle Innovation, dem die Agenda dieses Programms ein echtes Anliegen ist.

*Wien, im Juli 2023*

*Hannes Werthner, Ludger Hagedorn*

Hannes Werthner ist Mitinitiator der Digital-Humanism-Initiative und ehemaliger Dekan der Fakultät für Informatik der TU Wien (2016–2019); Ludger Hagedorn ist Permanent Fellow am Institut für die Wissenschaften vom Menschen (IWM).



## Pathologien

### Was kann schon schiefgehen?

1951 startete der BBC Home Service eine Rundfunksendung mit dem Titel *The '51 Society*, in der jeweils ein renommierter Experte einen Vortrag hielt, gefolgt von einer Podiumsdiskussion. Alan Turing wurde dreimal eingeladen. Beim ersten Mal hielt er einen Vortrag mit dem Titel »Intelligente Maschinen, Eine ketzerische Theorie«. Darin stellte er die These auf, dass

»sich Maschinen konstruieren lassen, die das Verhalten des menschlichen Geistes sehr genau simulieren. Sie werden manchmal Fehler machen, und manchmal werden sie neue und sehr interessante Aussagen treffen, und im Großen und Ganzen wird ihr Output ebenso beachtenswert sein wie der Output eines menschlichen Geistes. ... Nehmen wir an, dass es tatsächlich möglich wäre, diese Maschinen zu erzeugen, und überlegen uns, welche Folgen ihre Konstruktion hätte. Ein solches Unterfangen würde natürlich auf großen Widerstand stoßen, es sei denn, wir hätten seit Galileis Zeiten große Fortschritte in Sachen religiöser Toleranz gemacht. Der Widerstand der Intellektuellen, die Angst haben, arbeitslos zu werden, wäre groß. Aber wahrscheinlich würden sich die Intellektuellen irren. Es gäbe viel zu tun, um zu verstehen, was die Maschinen sagen wollen, d. h. um die eigene Intelligenz auf dem von den Maschinen gesetzten Standard zu halten, denn es scheint wahrscheinlich, dass, wenn die maschinelle Denkmethode einmal eingesetzt hat, es nicht lange dauern wird, bis