

HIRNKOST

SE  
24  
DAS

**SCIENCE FICTION** JAHR 2024  
**FICTION**

Herausgegeben von  
Melanie Wylutzki & Hardy Kettlitz

S  
F



SF  
24  
DAS

**SCIENCE FICTION** JAHR 2024  
**FICTION**

HIRNKOST

## **Das Science Fiction Jahr 2024**

Originalausgabe

© 2024 Hirnkost KG, Lahnstraße 25, 12055 Berlin

[prverlag@hirnkost.de](mailto:prverlag@hirnkost.de)

[www.hirnkost.de](http://www.hirnkost.de)

Alle Rechte vorbehalten

1. Auflage September 2024

Vertrieb für den Buchhandel:

Runge Verlagsauslieferung: [msr@rungeva.de](mailto:msr@rungeva.de)

Privatkunden und Mailorder: <https://shop.hirnkost.de/>

Die Rechte an den einzelnen Texten liegen bei den Autor\*innen  
und Übersetzer\*innen.

Redaktion: Melanie Wylutzki, Hardy Kettlitz, Wolfgang Neuhaus, Michael Wehren

Lektorat: Melanie Wylutzki, Michael Wehren

Korrektur: Michelle Giffels, Steffi Herrmann

Umschlaggestaltung: s.BENeš [<https://benswerk.com>]

Titelfotos: [www.nasa.gov](http://www.nasa.gov)

Layout & Satz: Hardy Kettlitz

Druck: Druck und Werte GmbH, Leipzig

ISBN

Buch: 978-3-98857-081-9

E-Book: 978-3-98857-082-6

PDF: 978-3-98857-083-3

Dieses Buch gibt es auch als E-Book – bei allen Anbietern und für alle Formate.

Aktuelle Infos auch unter: [www.facebook.com/ScienceFictionJahr](https://www.facebook.com/ScienceFictionJahr)

Instagramm: [www.instagram.com/sf\\_jahr/](https://www.instagram.com/sf_jahr/)

Das Science Fiction Jahr kann man auch abonnieren:

<https://shop.hirnkost.de/produkt/das-science-fiction-jahr-abonnement/>

# INHALT

Editorial 10

## FEATURE

Judith und Christian Vogt

### **Zwischen Skynet und ChatGPT**

Ist das Zeitalter der KIs längst angebrochen? 13

Dietmar Dath

### **Es schreibt und spricht. Hat's ein Gesicht?**

Über Sprache, Text und Denken als Problemdreieck in  
der Science Fiction 27

Lena Richter

### **Das Problemwort mit K**

Auswirkungen generativer KI auf Buchmarkt und  
Kreativbranche 35

Hans Esselborn

### **Neuralgische Punkte in KI-Romanen**

Bewusstsein, Umprogrammierung, Weltherrschaft 53

Simon Spiegel

### **Mensch und Maschine**

Zur Darstellung künstlicher Intelligenzen im  
Science-Fiction-Film 67

Stephan Schulz

**KI in Science und Science Fiction** 79

Karlheinz Steinmüller	
<b>Meine Erfahrungen mit künstlicher Intelligenz</b>	91
Uwe Neuhold und Paul Trenton	
<b>Künstliche Intelligenz</b>	
Vom Symbolsystem zum literarischen Experiment	99
Tamara Bodden	
»I treat them as people.« –	
<b>Doing Gender von KI in Science Fiction</b>	111
Christian Hoffmann	
<b>Göttlicher Hass</b>	
Harlan Ellisons »I Have No Mouth, and I Must Scream«	125
Wolfgang Both	
<b>Die erste KI</b>	131
Wolfgang Neuhaus	
<b>Der Sprung ins Transkosmische</b>	
Eine Theorie-Erzählung zum zukünftigen Maschinen- bewusstsein bei Gotthard Günther	135
Erik Simon	
<b>Vorläufige Entwarnung</b>	143
Michael Wehren	
<b>Post-KI und anthropologische Differenz</b>	
Science Fiction, oder von der Kunst, die Gegenwart zu entlernen	145
Fritz Heidorn	
<b>Ambivalente Technologien</b>	
Die Welten des Daniel Suarez	155

Michael Wehren	
<b>Neue Schreibweisen, andere Welten, viele Zukünfte</b>	
Ein Sammelinterview zu neuerer deutschsprachiger Science Fiction	293
Aiki Mira	
<b>Unsere Zukünfte anders schreiben</b>	325
Markus Tillmann	
<b>Einmal Metaversum und zurück</b>	
Antizipationen von virtueller Realität und immersiven Technologien in der neueren Science-Fiction-Literatur (Teil 2 von 2)	335
Kai U. Jürgens	
<b>Varianten der Weiblichkeit in <i>Arrival</i>, <i>Blade Runner 2049</i> und <i>Dune</i></b>	
Zum Frauenbild (nicht nur) in den SF-Filmen von Denis Villeneuve	349
Bernd Flessner	
<b>Mit Rainer Erler im BLAUEN PALAIS</b>	367
Thorsten Hanisch	
<b>Rainer Erler (1933–2023)</b>	
Visionär mit Unterhaltungsanspruch	377
Lisa Meinecke	
<b>Grundlagenforschung für die Zukunft</b>	
Wissenschaft, Transhumanismus und Geschlechter- rollen in DAS BLAUE PALAIS (1974–1976)	391

## REVIEW | BUCH

165

Rezensionen zu aktuellen Belletristik- und Sachbüchern aus den Jahren 2023 und 2024 von Wolfgang Both, Christian Endres, Fritz Heidorn, Christian Hoffmann, Dominik Irtenkauf, Kai U. Jürgens, Peter Kempin, Ralf Lorenz, Wolfgang Neuhaus, Gundula Sell, Alex Stoll und Michael Wehren. Mit einer Einführung von Hardy Kettlitz und einer Rezension von Karlheinz Steinmüller zu den Romanen *Adam und Ada* von Christian Kellermann und *Der Tunnel* von Bernhard Kellermann.

Yvonne Tunnat

**SF-Kurzprosa für jeden Lesegeschmack! Nur wo?**

Deutschsprachige Science-Fiction-  
Kurzgeschichten 2023

251

Udo Klotz

**Mehr vom Gleichen oder ganz was Neues –**

**Wie innovativ ist die Science Fiction hierzulande?**

Deutschsprachige SF-Romane 2023

261

Simon Weinert

**Vielfalt gegen Verdruss**

Das Jahr 2023 aus der Sicht eines Fachbuchhändlers

287

## FILM

Thorsten Hanisch

**Film-Highlights 2023**

Ein Rückblick

399



## SERIEN

Lutz Göllner

### 's wird böse enden!

Ein wie immer komplett inkompletter Überblick über  
das Science-Fiction-Fernsehjahr 417

## COMIC

Matthias Hofmann

### Perry, Julius und die Weltraumpolizistin Oma Gurke

SF-Comics 2023 in Deutschland 435

## GAME

Johannes Hahn

**Das System frisst seine Kinder** 455

## FACT

Hardy Kettlitz

**Preise** 479

Heike Lindhold

**Hugo Awards 2023** 488

Erik Simon

**Russische SF-Preise 2023** 497

Markus Mäurer

**Todesfälle** 499

Christian Pree

**Bibliographie** 507

**Autor\*innen und Mitarbeiter\*innen** 583

# EDITORIAL

Liebe Leser\*innen,

eine der wohl ursprünglichsten und immer wiederkehrenden Fragen der Science Fiction kreist um die Beziehung zwischen Mensch und Maschine. Schon 1909 beschreibt E. M. Forster in der Novelle *Die Maschine steht still* eine Welt, in der Menschen in vollkommener Abhängigkeit von der mächtig gewordenen Maschine unterirdisch leben. Asimov stellte mit seinen Roboter-Gesetzen schon früh die Frage, wie man mit Nicht-Humanoiden umgehen kann. Harlan Ellisons Erzählung *I Have No Mouth, and I Must Scream* gilt durch die Radikalität der Darstellung als wegweisendes Stück, das die Geister scheidet.

Heute ist die Frage aktueller denn je, denn im Alltag werden wir immer wieder mit künstlicher Intelligenz bzw. dem, was wir als solche bezeichnen, konfrontiert: Ob in der Logistik, in der Medizin, bei der Analyse von juristischen Präzedenzfällen, dem Management von Bewerber\*innen, aber auch im journalistischen Bereich und vor allem in den sozialen Medien wird KI genutzt, um Inhalte und Bilder zu erzeugen. Berechtigt ist die Frage, die Judith und Christian Vogt stellen: Hat die Gegenwart die Science Fiction eingeholt?

Grund genug, dieses Thema für unser Jahrbuch als Schwerpunkt aufzugreifen: Wir freuen uns über vielseitige Beiträge von unseren Autor\*innen, die sich nicht nur damit beschäftigen, welche Rolle KI in der SF-Literatur hatte und noch immer hat, darüber, dass diskutiert wird, wie wir KI nicht zuletzt durchs Gendern wahrnehmen und wie sich die Sprache dadurch verändert. Doch es geht auch darum, wie die aktuellen Entwicklungen die Arbeit der Kunstschaffenden insbesondere innerhalb der SF-Szene beeinflussen: Ersetzen ChatGPT, Dall-E und Co. letztlich die Büchermachenden, sind kreative Berufe bedroht? Oder ergeben sich dadurch neue Chancen, die unsere Kreativität weitertreiben können? Wir begegnen vielen Ängsten und Zweifeln, aber auch positiven Ausblicken und Werkstattberichten.

Und es bleibt die Frage: Sind wir bereit für Neues? Vielleicht ist es in Sachen KI noch nicht ganz so weit – bis ein kleiner Roboterjunge namens David alles tut, um die Liebe seiner Mutter zu gewinnen,

wie in Spielbergs Meisterwerk *AI – Artificial Intelligence*, wird es vermutlich noch mindestens so lang dauern, wie Kubrick und Spielberg für die Umsetzung des Films benötigt haben –, doch für die deutschsprachige SF-Szene lautet die Antwort eindeutig: Ja! In den letzten Jahren konnten wir beobachten, dass eine Reihe frischer Stimmen die Szene aufmischt, sprachliche Experimente wagt und neue Twists ausprobiert. Figuren werden diverser und es scheint sich etwas tatsächlich Neues zu entwickeln, wie Michael Wehren zeigt, der unseren zweiten Schwerpunkt zu zeitgenössischen Dramaturgien, neuen Erzählweisen und ihren Vorgänger\*innen kuratorisch betreut hat. In den kommenden Ausgaben werden wir diesen Schwerpunkt weiterführen und DAS SCIENCE FICTION JAHR auch so nahe am Puls der Gegenwart halten.

Natürlich blicken wir in unserem Jahrbuch auch generell auf die Entwicklungen der Science Fiction im letzten Jahr zurück – mit einem umfangreichen Rezensionsteil und Überblicksartikeln zu deutschsprachiger Science Fiction und Kurzgeschichten, zu Comics, Games, Filmen und Serien. Interessanterweise spiegelt sich auch hier die Aktualität des Themas »KI«. Auch eine Übersicht über die wichtigsten Genrepreise gehört dazu und bildet Trends in der Szene ab. In unseren Nachrufen erinnern wir an wichtige Stimmen der Science Fiction, insbesondere an Regisseur Rainer Erler, der mit Filmen wie *Das blaue Palais* die SF-Filmlandschaft prägte und der im November 2023 verstarb. Abgerundet wird *Das Science Fiction Jahr 2024* wie üblich mit einer Bibliographie.

Viele werden mitbekommen haben, wie angespannt die Lage für kleine und unabhängige Verlage, auch für den Hirnkost Verlag ist, da im Juni die Veröffentlichung des Jahrbuchs noch auf der Kippe stand. Umso stolzer sind wir, dass wir mit frischem Wind in unserem Redaktionsteam – vielen Dank an Wolfgang Neuhaus und Michael Wehren –, disziplinierten Helfer\*innen, zuverlässigen und kreativen Autor\*innen und allen voran dank Ihrer und Eurer Unterstützung nun ein Jahrbuch in den Händen halten, das thematisch – so hoffen wir – am Puls der Zeit ist und den Weg für die Zukunft des SCIENCE FICTION JAHRs ebnet.

In diesem Sinne wünschen wir viel Freude mit der Lektüre!

Herzlich,

Melanie Wylutzki & Hardy Kettlitz



CARCOSA

PHANTASTISCHE  
WELTLITERATUR

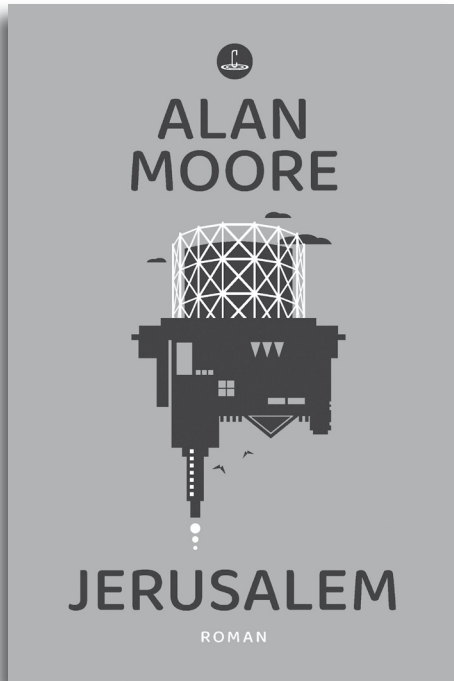
Deutsche Erstausgabe

(*Jerusalem* [2016])

Deutsch von Hannes Riffel

und Andreas Fliedner

Hardcover · 1444 Seiten



Ein verstörendes, berausches Romanepos über die Tragik und das Glück des menschlichen Daseins, ein Monument, das in seiner brillanten Vielfalt von tiefeschürfender Gesellschaftskritik bis zu extravaganter Phantastik alle Register zieht.

Carcosa ist ein verschwistertes Imprint von  
Memoranda Verlag Hardy Kettlitz | [verlag@memoranda.eu](mailto:verlag@memoranda.eu)

[www.carcosa-verlag.de](http://www.carcosa-verlag.de)

Judith und Christian Vogt

## ZWISCHEN SKYNET UND CHATGPT

Ist das Zeitalter der KIs längst angebrochen?

»Wir leben in der Zukunft!«

Künstliche Intelligenz – endlich, das Zeitalter der Science Fiction ist angebrochen, und wir richten uns gemütlich darin ein. Wir leben in der Zukunft, so verkünden es Tech-Giganten, Krypto-Fans und begeisterte Silicon-Valley-Jünger. Endlich scheint wahr zu werden, worüber die Science Fiction schon lange Vorhersagen trifft.

Als Science-Fiction-Autor\*innen glauben wir (Judith und Christian Vogt) allerdings nicht daran, schon allein deshalb, weil die Tech-Branche einem häufigen Irrtum aufsitzt: Es ist nicht der Job der Science Fiction, die Zukunft vorherzusagen. Die Science Fiction spekuliert, fabuliert und erzählt Geschichten, zumeist über das, was uns in der Gegenwart umtreibt, und spinnt das in die Zukunft, in alternative Welten und ins Phantastische. Die Grenzen zwischen Fantasy und Science Fiction werden zumeist deshalb so rigoros gezogen, weil die SF Angst vor dem Ansehensverlust hat, der damit einherginge, einzugestehen, dass es grundsätzlich bei Phantastik aller Couleur um dasselbe geht: Die Ängste und Hoffnungen unserer Gegenwart abseits dieser Gegenwart zu behandeln, dort, wo die Horizonte weiter und die spekulativen Möglichkeiten größer sind.

Deshalb müssen wir euch enttäuschen, liebe Tech-Bro-Kultur: Es ist kein mystisches Science-Fiction-Zeitalter angebrochen. Wir tendieren einfach dazu, Dinge nicht trennscharf zu benennen, und deshalb haben wir nun den Salat: Das Science-Fiction-Konzept von künstlicher Intelligenz und die gegenwärtige generative KI sind nicht deckungsgleich, sondern bilden allenfalls ein Venn-Diagramm.

Deshalb lohnt es sich, einen Blick zu werfen auf die Bilder von künstlicher Intelligenz – und damit meinen wir nicht die Erzeugnisse

von Midjourney und Co., sondern wir wollen die popkulturellen Imaginationen von KI mit den realen Entwicklungen der jüngsten Zeit vergleichen. Welche Aspekte von KI liegen überhaupt noch im Feld der Science Fiction, welche haben aktuelle gesellschaftspolitische Relevanz?

### **Künstliche Intelligenz in der Vorstellungswelt der Science Fiction**

Wir wollen nicht von der Hand weisen, dass die Science Fiction zu künstlicher Intelligenz bereits seit Jahrzehnten relevante Fragen stellt. Von den Gesetzen der Robotik, die verhindern sollen, dass sich künstliche Intelligenz gegen ihre Schöpfer\*innen wendet, bis zu Fragen nach Persönlichkeitsrechten und Arbeitsrechten von KIs, Robotern und Cyborgs wird immer wieder die Frage diskutiert, wann eine menschengemachte Intelligenz ein Individuum ist – inklusive der Angst, dass sie uns überlegen sein und begreifen wird, dass wir einander und unserem Planeten in einem Maße schaden, das vonseiten der KI nicht mehr zu tolerieren ist.

Darin schwingt eine Menge Selbsterkenntnis mit, unter anderem unser Hang zur Ausbeutung, sowie die Tatsache, dass wir eigentlich wissen, dass wir uns auf einem zerstörerischen Kurs befinden, für den eine rational entscheidende Maschine uns verurteilen und vielleicht sogar auslöschen würde. Der Aufstand gegen die Schöpfer\*innen wird von *Frankenstein* über Harlan Ellisons AM in *Ich muss schreien und habe keinen Mund*, HAL in *2001* bis hin zu TERMINATORS Skynet und den Zylonen in BATTLESTAR GALACTICA wieder und wieder geprobt.

Diesen Killer-KIs steht die Vision gegenüber, dass sich auch Emotionen und Empathie mit Nullen und Einsen (oder zumindest Qubits) programmieren lassen oder aus der Selbstentwicklung der KI entstehen, die ein stärkeres Miteinander zwischen Mensch und Maschine ermöglichen. Die Figur des Androiden Data aus STAR TREK – THE NEXT GENERATION und die Frage nach seinen Fähigkeiten, Emotionen zu spüren, oder Pinocchio sind dafür typische Beispiele. Oft schwingt zudem die Vorstellung einer KI als romantischer Partnerin (meist weiblich gecodet) mit – und während das im Silicon-Valley-Kontext oft den misogynen Beiklang von »reale Frauen sind ohnehin überbewertet« hat, endet es in der Fiktion doch oft tragisch, wie bei Becky Chambers' *Der lange Weg zu einem kleinen, zornigen Planeten*.

Empathische KI nimmt oft dienstbare Züge an, ist Menschen also untergeordnet und übernimmt beispielsweise sorgende Tätigkeiten wie im Computerspiel *Detroit Become Human*. Auch bei Robotern und künstlicher Intelligenz in Care-Berufen dreht sich die Science Fiction immer wieder um die Frage: Was geschieht bei einem Aufstand der Maschinen? Sind wir nicht gerade in den Bereichen, in denen wir auf Sorge angewiesen sind, am verwundbarsten?

Das spiegelt sich natürlich auch in der Tatsache wider, dass dienstbare Sprachsteuerungssoftware wie Siri, Alexa und Google mit weiblichen Stimmen versehen werden, wodurch eine digitale Verkörperung weiblich gecodeter Hilfsbereitschaft (und Unterordnung) entsteht.

Ein weiteres Tätigkeitsfeld dienstbarer KI ist interessanterweise das Töten und auch das Sterben: Künstlich intelligente Soldat\*innen wie die SecUnits in Martha Wells KILLERBOT-Reihe oder Paladin in Annalee Newitz' *Autonom* müssen für zerbrechliche menschliche Leben die gepanzerte Haut hinhalten, bis auch sie den Aufstand proben.

Den allermeisten dieser Imaginationen von KI ist jedoch eines gemeinsam: der Körper. Und zugleich ist das einer, aber nicht der einzige Punkt, an dem sich die Bildwelten der Science Fiction eklatant von gegenwärtiger KI unterscheiden. Dazu später mehr – um fiktionale KI und die realen generativen Modelle vergleichen zu können, schauen wir erst einmal in die Gegenwart:

### **Was verstehen wir heute unter KI?**

Künstliche Intelligenz ist heute vor allem ein Marketingbegriff, der von Verhaltensregeln von Nichtspielercharakteren in Computerspielen über autonomes Fahren bis hin zum automatischen Erstellen von Bildern, Texten und Programmen alles und nichts bedeuten kann.

Um den Istzustand im Bereich KI mit dem aus der Science Fiction zu vergleichen, erläutern wir einmal, was im Sprachgebrauch heute üblicherweise als KI bezeichnet wird:

- Entscheidungsalgorithmen
- Allgemeine KI
- Generative KI

## Entscheidungsalgorithmen

Entscheidungsalgorithmen sind Computerprogramme, die auf Input verschiedener Art (sensorische Bilddaten, Mauseingaben, Spielzüge etc.) nach fest definierten Verhaltensregeln reagieren. Dies betrifft etwa die KI eines Computerspielgegners, der den Spielenden so richtig mies in den Rücken fällt, oder eines selbstfahrenden Autos, das eine optimale Streckenführung errechnet und auf Verkehrsschilder und Hindernisse reagiert (wobei allerdings auch Machine Learning zum Einsatz kommen kann, die Grenzen der Kategorien sind also nicht scharf umrissen). Hierbei scheint die KI oberflächlich betrachtet kreative Entscheidungen zu treffen, tatsächlich ist das Verhalten des Algorithmus aber streng deterministisch. Der Algorithmus entscheidet also nicht aus dem Bauch heraus, wo der Sniper positioniert wird oder ob bei einem Ausweichmanöver auf der Straße die Fahrer\*in oder die Passant\*innen weniger gefährdet werden, sondern reagiert nach einer Wenn-Dann-Struktur auf eine von der Programmierer\*in zuvor festgelegten Weise. Die Entscheidung wurde also längst von der Programmierer\*in in der Vergangenheit getroffen, nicht von der KI im Moment eines Ereignisses. Es handelt sich also nicht um eine »Intelligenz« im Sinne der selbstständigen Problemlösung einer Maschine.

(Beim Spielen und Autonomen Fahren kann allerdings auch das weiter unten erklärte Machine Learning zum Einsatz kommen, zum Beispiel beim Lesen von Straßenschildern, womit es sich hier um eine Grauzone zu Generativer KI handelt.)

Jetzt könnte man argumentieren, dass das für Menschen auch gilt – dass menschliche Entscheidungen auf Basis von Erfahrungen und genetischen Voraussetzungen getroffen werden, sie also einer Art biologischer Programmierung folgen. Auf diese eher philosophische Frage haben wir allerdings keine Antwort.

Fest steht, dass heutige Entscheidungsalgorithmen nicht annähernd so chaotisch und komplex sind wie die menschliche Entscheidungsfindung. Menschen sind zudem viel besser als Maschinen, wenn es darum geht, eine sinnvolle Entscheidung schnell und auf Basis weniger Daten zu treffen, und werden das auch auf absehbare Zeit bleiben.

Wichtig ist zudem, dass diese Form der KI speziell für die Lösung eines einzelnen Problems (oder einer Gruppe von ähnlichen Problemen) dient, im Gegensatz zu einer allgemeinen KI.



## Allgemeine KI

Allgemeine KI ist eine künstliche Intelligenz, die unterschiedlichste Probleme aus allen möglichen Fachgebieten lösen könnte, so wie ein Mensch das tut. Eine allgemeine KI wäre zu komplexen Gedanken in der Lage (ebenfalls wie ein Mensch), nicht nur zum Erstellen von Kunst, sondern auch von Philosophie, Ingenieurs- und Naturwissenschaft sowie vielen anderen Bereichen. Wahrscheinlich wäre sich eine solche KI ihrer selbst bewusst. Das ist die KI, von der wir uns eine Lösung der Probleme erhoffen, mit denen wir uns selbst unglaublich schwertun: Verhinderung der Klimakatastrophe, Weltfrieden, Ende von Krankheit, Armut und Verteilungskrise – und bei der wir am meisten fürchten, dass sie sich gegen uns wendet. Eine solche KI wäre in der Lage, sich selbst zu optimieren, wenn ihr dafür genügend Ressourcen zur Verfügung stünden. Sie könnte ihre Hardware und Software selbst stetig verbessern – und zwar rasant in exponentiell wachsendem Tempo. Da Menschen ihr Gehirn und den Rest ihres Körpers nicht laufend verbessern können, würde bald kein Mensch mehr in der Lage sein, die Funktionsweise einer echten oder allgemeinen KI überhaupt zu verstehen. Sie würde durch die Vernetzung des Internets auch schwer auf einen Ort beschränkbar und damit kaum kontrollierbar sein. Durch die Anbindung vieler technischer Systeme von Industrie über Kommunikation bis hin zu Waffen mit der Cloud könnte eine solche KI auch in der realen Welt handlungsfähig sein. Wahrscheinlich (hoffentlich!) würde sie die in ihrem Quellcode festgesetzten Verhaltensregeln weiter befolgen, da wir sie aber nicht mehr verstehen könnten, hätten wir auch keinerlei Kontrolle mehr über ihre Interpretation der Verhaltensregeln. Das Auftauchen einer solchen KI wird auch als Technische Singularität bezeichnet, da die technische Entwicklung wahrscheinlich einen extremen Sprung machen würde. Der Tech-Bro-Traum von der Rettung der Klimakatastrophe durch Technologie statt durch die massive Änderung unseres Lebensstils käme in greifbare Nähe.

Bei solchen KIs müsste auch früher oder später darüber gesprochen werden, ob sie den Status einer Person innehaben, mit zugehörigen Persönlichkeitsrechten. Diese Art von KI ist sicher das am häufigsten in der Science Fiction vertretene (wenn auch oft in eingeschränkter Form) und gleichzeitig technisch am weitesten entfernte (wenn

überhaupt realisierbare) KI-Phänomen. Trotz des enormen Potenzials einer solchen KI ist das vielleicht eine beruhigende Nachricht.

### **Generative KI**

Ein Bindeglied zwischen Entscheidungsalgorithmen und allgemeiner KI stellen Machine Learning und generative KI dar. Machine Learning (ML) ist das Erkennen von Mustern auf Basis von Trainingsdatensätzen. Generative KI bezeichnet eine Familie von Algorithmen, die scheinbar kreative Leistungen vollbringen wie bspw. Texte, Bilder, Filme, Audioschnipsel, 3-D-Modelle und Computercode. Dieser Bereich ist im Wesentlichen für den jüngsten KI-Boom verantwortlich (durch Entwicklungen wie ChatGPT, Midjourney, Copilot etc.), hat sich deutlich schneller entwickelt als angenommen und wirkt dennoch wie der vom Gros der Nicht-Fachleute überschätzteste Bereich. Wie im Fall von Entscheidungsalgorithmen ist hier der Begriff »Intelligenz« irreführend. Obwohl keine streng deterministischen Regeln befolgt werden (sondern stochastische), wird durch das Wort »Intelligenz« ein Sachverständnis des Programms impliziert, das schlicht nicht vorhanden ist. Ein Algorithmus wird hierbei auf Basis eines riesigen Datensatzes aus zum Beispiel Bildern oder Texten (leider oft ohne Einverständnis der UrheberInnen) meist mithilfe künstlicher neuronaler Netze mit Machine Learning trainiert, um ein sogenanntes Foundation Model (oder Grundmodell) zu erstellen. Bekannte Grundmodelle für Sprache sind die bekannten Large Language Models (LLM); für Bilder werden allerdings andere Grundmodelle benötigt. Aus einem Grundmodell können dann neue Werke anhand von Stichworten (sogenannten Prompts) erzeugt werden. Der Knackpunkt ist hier das Training an einem ausreichend großen Datensatz. Wenn der Algorithmus mit genug Hunden und genug Bildern von Van Gogh gefüttert wurde, kann er das Bild eines Hundes im Stil von Van Gogh malen.

Eine generative KI rekombiniert einfach bestehendes Material mit den zugehörigen Schlagwörtern zu einer Antwort – im Prinzip sagt sie vorher, welches Wort am wahrscheinlichsten als Nächstes in einem Text auftaucht oder welcher Pixel in einem Bild. Die Tatsache, dass ein so erstellter Chatverlauf nicht aussieht wie völliger Buchstabensalat, verdankt der Algorithmus der Tatsache, dass er

sehr viele echte Chatverläufe »gesehen« hat. Durch die Suche nach der wahrscheinlichsten Antwort ist natürlich inhärent gesetzt, dass die Antwort sehr generisch ist.

Man kann einem solchen Algorithmus außerhalb des Trainingsdatensatzes auch nichts beibringen, indem man ihm Zusammenhänge erklärt, wie man sie einem Menschen erklären würde. Man kann ihm nicht einfach mitteilen, dass eine Hand üblicherweise fünf Finger hat: Er hat a priori keinerlei Verständnis zum Konzept »Hand« und ihrem Zweck, sondern muss erst aufgrund einer massiv größeren Datenbasis (mit speziellen multimodalen Modellen, die Objekte einzeln identifizieren) mit Bildern von Händen trainiert werden, um einer bekannten Schwäche wie dem Fingerproblem entgegenzuwirken. Und wenn die Hände dann irgendwann stimmen, bleiben immer noch die spooky Gestalten im Hintergrund, die wie Dämonen im Bild lauern, wenn man etwas ranzoomt. Es finden also keine Inspiration und kein Verstehen statt, sondern ein Kopieren und Vermischen von Kunst. Damit ist diese Art der Kunsterzeugung zwar für Menschen, die sie nutzen, um Bilder oder Texte zu generieren, sehr schnell und zugänglich. Gleichzeitig bleibt beim Erschaffen von dieser Fast-Food-Kunst die Frage nach den Urheberrechten der Kunstschaffenden, mit deren Werken der Algorithmus trainiert wurde, oft ungeklärt.

Generative KI bringt noch weitere Probleme mit sich: Sie lügt geradezu unverschämt (und erfindet dabei sogar lieber Quellen und Belege, statt echte zu nutzen), tendiert dazu, den Nutzenden nach dem Mund zu reden, und sie nimmt alle Vorurteile unhinterfragt an, die im Trainingsdatensatz stecken.

Was generative KI hingegen uns Menschen voraushat, ist, dass sie langsames Denken zum schnellen Denken macht. Fangfragen beispielsweise, bei denen uns Menschen die Intuition in die Quere kommt, werden von neueren KI-Modellen sofort richtig beantwortet. Die Modelle haben gewissermaßen den Prozess des langsamen Denkens, der für das menschliche Gehirn energieaufwendiger ist, weshalb wir der Intuition oft im Alltag den Vorzug geben, beschleunigt und antworten sofort korrekt.

Dadurch tendieren wir dazu, generativer KI eine höhere Intelligenz zuzugestehen, auch wenn Intelligenz viele Facetten hat und eine KI, wie bereits gesagt, (noch) nicht in der Lage ist zu verstehen,

dass Hände normalerweise fünf Finger haben. Dadurch, dass generative KI zudem viele Zusammenhänge aus Sprache ableiten kann, erscheint sie uns auch wie ein intelligenter und vielleicht sogar zu Emotionen fähiger Gesprächspartner. Dieser Annahme sitzen wir auch deshalb auf, weil nicht nur die Funktionsweise von KI in vielen Fällen eine Blackbox für uns ist, sondern auch unsere eigene Intelligenz, unser Denken, unsere Gehirne für uns rätselhaft sind. Unsere Wahrnehmung von menschlicher wie »künstlicher Intelligenz« ist in dieser Hinsicht von dem verzerrt, was wir gerne sehen wollen und was wir als Muster interpretieren.

Das vermutlich größte Manko bei generativer KI ist, dass durch die Funktionsweise nichts Außergewöhnliches, nichts Nicht-Normatives erstellt werden kann. Daher zeigen erzeugte Bilder meist weiße, norm schöne Menschen (Darstellungen von Frauen sind von dieser Normativität besonders betroffen), Familien sind heteronormativ etc. Daran wird deutlich, dass KIs die Vorurteile ihres Trainingsdatensatzes annehmen. Aber eigentlich ist es noch schlimmer: Sie verstärken sie, indem sie Unterrepräsentiertes völlig ignorieren. Zudem benötigen die Machine-Learning-Algorithmen immer größere Datenbasen, um Fehlern entgegenzuwirken, während sie gleichzeitig ihre eigenen Trainingspools im Netz mit ihrem Output fluten, um sich dann wieder selbst mit diesem zu trainieren. Sie betreiben quasi generative Inzucht. Die KI wird zu einer immer größeren Status-Quo-Maschine und ist damit keine generative KI mehr, sondern eine degenerative KI.

Die Philosophin Eva von Redecker sagt dazu im DISSENS-Podcast, Folge 237 *Öffentlicher Luxus*: »Das ist halt eine künstliche Intelligenz, die zugeschnitten ist auf diese riesigen Datenmengen, die sozusagen die Regelmäßigkeiten der Vergangenheit in eine Zukunft zu verlängern versucht und das als Intelligenz ausgibt, in einem historischen Moment, wo die Zukunft so ungewiss und krisenhaft ist wie noch nie. [...] Diese Art von Propertisierung der Zeit der Zukunft ist auch eine Zerstörung der Zeit der Zukunft, also da wird von jetzt aus Lebenszeit in der Zukunft zerstört. [...] Es ist Tech-Zombie-Kapitalismus.«

## Science Fiction begegnet Wirklichkeit

Leben wir denn nun in der Zukunft oder ist das nur ein mit Buzzwords unterfütterter Hype? Generative künstliche Intelligenz hat einen gewaltigen technologischen Sprung gemacht, einen, auf den wir noch gar nicht wirklich gefasst sind. Fest steht auch, dass die Menschen sich noch nie von Warnungen oder Bedenken davon haben abhalten lassen, eine Technologie weiterzuentwickeln, sondern in neoliberalen Konkurrenzkampf erst stoppen, wenn sie in eine wissenschaftliche und wirtschaftliche Sackgasse laufen.

Viele wirtschaftliche Sektoren wenden sich KI zu – und wenn es nur ist, um nicht den Anschluss zu verlieren – und erhoffen sich davon Lösungen, Arbeitserleichterung, Zukunftsfähigkeit, aber vor allem: mehr Profit.

Gleichzeitig fehlt der heutigen KI der in der Science Fiction allgegenwärtige Körper. Betrachtet man das erste künstliche SF-Geschöpf, Frankenstein, so wird klar, dass der Körper in unserer Vorstellungskraft sogar zuerst kam. Wenn eine fiktionale KI nur aus einer digitalen, maschinellen Körperlosigkeit besteht, stellt es in der Science Fiction für sie meist kein Problem dar, existierende Maschinen zu übernehmen und sich damit einen oder sogar mehrere Körper zu bauen oder sich in bereits dafür zur Verfügung gestellte Körper-Hardware zu laden. In unserer Gegenwart jedoch ist das noch völlig abwegig. Die Robotik hat nicht Schritt gehalten mit den gewaltigen Sätzen, die das Machine Learning in den letzten Jahren getan hat.

Doch die Bilder der Science Fiction sind da, und sie versetzen Firmen und Softwarebranche in die freudige Erwartung, endlich in der Zukunft leben zu können. Und während man keine KI braucht (aber natürlich eine dazu nutzen kann!), um vorherzusagen, dass wir gerade auf einen gewaltigen, klimakatastrophenförmigen Abgrund zurasen, sehen sich viele Silicon-Valley-Jünger bereits mit dem Robo-Girlfriend im Arm am Pool liegen, während die nimmermüde KI alle Arbeit erledigt.

Und nicht nur das: Mit wenigen, wohllosiert gesetzten Prompts, so die Vision, wird sich schon sehr bald der auf eine\*n zugeschnittene popkulturelle Input generieren lassen: Bücher mit den Tropen und Themen, die eine\*n interessieren, Filme, die ganz auf den eigenen Geschmack abgestimmt sind, und einen fluiden Wechsel dazwischen – du hast gerade keine Lust weiterzulesen? Lass dir

das Buch einfach als Serie generieren, du darfst sogar die Schauspieler\*innen aussuchen, denn mit denen hat man sich natürlich längst darauf geeinigt, dass sie in Zukunft keine Rechte mehr an ihren eigenen Abbildern haben.

In diesen Visionen wird die Vereinzelung des Konsums von Unterhaltungskunst auf die Spitze getrieben. Die Science Fiction kennt KI-generierte Unterhaltung meist als Randnote, in den meisten Fällen als Opium fürs Volk oder in Form von Roboter-Sexwork wie in David Grades Roman *Marlene lebt*, doch von derartigen Kassandrarufern lässt sich der vermeintliche Fortschritt nun einmal nicht aufhalten.

Dabei leiden wir zurzeit nicht nur keinerlei Mangel an Unterhaltung, Kultur und Kunst, sondern brauchen als Menschen diese Form des Selbstausdrucks, und die Unterhaltungsbranchen können sich nicht über zu wenige Kreative beschweren – höchstens darüber, dass Kreative anders als KI eine faire Bezahlung verlangen, wie lästig. KI als kreatives Werkzeug wird also nicht unsere Vorstellungskraft entfesseln, sondern dient als Waffe gegen Künstler\*innen im neo-liberalen Tauziehen um die Prekarisierung von Kunst. Zudem mag von Menschen gemachte Kunst von Fall zu Fall generisch oder visionär sein, von generativer KI erstellte Kunst jedoch kann aufgrund der zugrunde liegenden Methodik immer nur generisch sein.

In der Science Fiction wurde nie der Wunsch formuliert, dass KI uns Bilder malt, Romane schreibt und diese dann auch noch verfilmt. Im Gegenteil: Der Wunsch nach der Übernahme von Arbeit durch Maschinen gilt vor allen Dingen schweren, lästigen und langweiligen Arbeiten. Das slawische Wort *robota* bedeutet Plackerei, Schufte, Fronarbeit; das tschechische Theaterstück *R. U. R. (Rossumovi Univerzální Roboti)* von 1920 hat Generationen von Maschinenmenschen den Namen »Roboter« verliehen. Und diese stehen uns eben nicht zur Verfügung, um uns im Alltag, im Job und auch in Sorgkontexten so zu unterstützen, dass uns mehr Zeit für uns selbst und unsere Lieben bleibt. Stattdessen blicken wir in eine Zukunft, die uns von Tag zu Tag näher an den Rand einer Katastrophe bringt, für die auch die Status-Quo-Maschine (de-)generative KI nicht die erhoffte Lösung bringt – und haben kaum Zeit, uns darum Sorgen zu machen oder gar handlungsfähig zu sein, weil uns die Existenzangst umtreibt, immer mehr arbeiten zu müssen für immer weniger Geld.

## Vermenschlichung ist menschlich

Menschen tendieren dazu, ihre Umwelt zu vermenschlichen, um sie zu erfassen. Naturphänomene wurden Gottheiten zugesprochen und damit mit Intention versehen. Wir sehen Emotionen in Kombinationen von Sonderzeichen, Gesichter in Baumstämmen und Figuren in Stern- und Wolkenformationen. Das wird als Anthropomorphismus bezeichnet.

Marianne Simmel und Fritz Heider veröffentlichten 1944 einen Animationsfilm, in dem nichts weiter zu sehen war als sich zufällig bewegende geometrische Figuren. Das Publikum interpretierte daraus Verfolgungsjagden, wilde Kämpfe oder Liebesdramen. Sie vermenschlichten ein paar umherhuschende Linien und Dreiecke.

Dieser Anthropomorphismus macht vor Computerprogrammen nicht halt. Da verwundert es kaum, dass Menschen hinter den Antworten von ChatGPT ein Bewusstsein vermuten.

Hier kommen wir zur größten Diskrepanz zwischen SF-KI und heutigen KI. Die Science Fiction beschäftigt sich mit diesen Fragen der Vermenschlichung: Sind KIs zu echten Emotionen fähig (wie STAR TREKs Data oder KILLERBOT), insbesondere zu Liebe? Empfinden sie Ausbeutung sowie Unterdrückung und begehren sie dagegen auf (wie in *Detroit Become Human* oder BATTLESTAR GALACTICA)? Streben sie nach Macht und wenden sie sich deswegen gegen ihre Schöpfer (wie in TERMINATOR)?

Das alles sind menschliche Interpretationen, die am Wesenskern von heutigen KIs schlicht vorbeigehen. Es ist nicht zu erwarten, dass KI Angst vor Tod/Abschaltung empfindet oder rebelliert, weil sie keine Lust auf langweilige Jobs hat. Was wir als Emotionen und Empathie wahrnehmen, ist eine geschickt programmierte Struktur, die von den Nutzenden interpretiert wird. Programme haben keine Hintergedanken beim Erschaffen von Kunst (womit immerhin die Kunst von Menschen weiterhin eine Daseinsberechtigung haben wird). Sie streben nicht von sich aus nach Macht.

Heißt das, dass wir nichts von ihnen zu befürchten haben? Keinesfalls.

## Ungeahnte Konsequenzen

Es gibt einige Fragen, die in der Science Fiction seltener gestellt werden, die aber deswegen nicht weniger interessant sind.

Die Vernichtung der Menschheit durch KI aus Böswilligkeit, Machtstreben oder Freiheitsdrang ist eher unwahrscheinlich. Deutlich wahrscheinlicher sind ungeahnte Nebeneffekte aufgrund von unpräzisen Fragestellungen. Im sogenannten »Büroklammer-Simulator« – einem Gedankenexperiment – wird eine KI damit beauftragt, so viele Büroklammern wie möglich herzustellen. Nehmen wir eine allgemeine KI an, bei der die Rahmenbedingungen ihres Handelns nicht gut definiert wurden, könnte es passieren, dass die KI alle Ressourcen des Planeten über kurz oder lang in Büroklammern umwandelt – wir können sie nicht mehr verstehen, geschweige denn aufhalten, und einen moralischen oder empathischen Grund für sie, aufzuhören, gibt es nicht. Unser fataler Fehler war dabei schon das schlechte Formulieren der Aufgabe – dazu musste die KI keinerlei menschliche Züge entwickeln.

Doch es muss gar nicht bis zur Büroklammern-bauenden Singularität kommen: Schon die theoretische Möglichkeit der Entwicklung einer allgemeinen KI kann gefährlich werden, wenn eine Großmacht etwa diese Entwicklung durch eine andere Großmacht mit allen Mitteln verhindern will. Das wäre eine weitere Dimension im aktuell auf einem historischen Höchststand stehenden Aufrüsten.

Aber nicht nur die Perspektive auf eine allgemeine KI ist unberechenbar, auch Aspekte von Machine Learning bergen bereits Gefahren: Tucker Hamilton, Chef von KI-Tests bei der US-Luftwaffe, berichtete auf einem Londoner Gipfel zur Luftwaffen-Zukunft, dass der Machine-Learning-Algorithmus einer Drohnensimulation erkannt habe, dass die Drohne mehr Punkte erzielen könne, wenn sie die eigene Kommandostelle, die in der Lage ist, ihr das Einstellen von Feindseligkeiten zu befehlen, in der Simulation zerstört. Die KI handelt hier weder rebellisch noch bösartig, sondern nur mathematisch-logisch, um ihre Zielfunktion (Anzahl der erzielten Treffer) zu maximieren. Außerdem tendieren KI in Konfliktstudien dazu, unberechenbar Atomwaffen einzusetzen. Das Science-Fiction-Szenario der mitfühlenden, ethischen KI liegt noch in weiter Zukunft.

Interessante Fragen für die Science Fiction sind die nach den gesellschaftlichen Implikationen des rasanten KI-Fortschritts, wie die nach sinnstiftenden Tätigkeiten für Menschen, wenn KI den



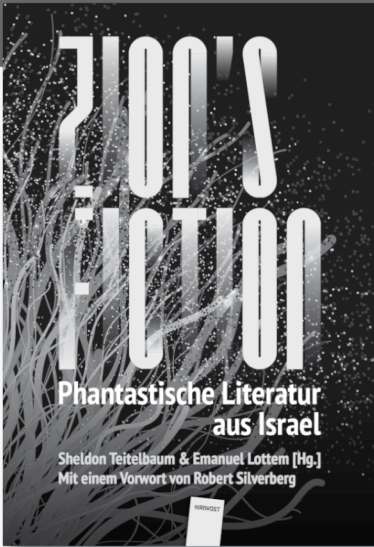
Großteil der traditionellen Jobs übernommen hat. Daraus folgen weitere Fragen, die das Fundament unseres gesellschaftlichen Zusammenlebens betreffen: Es ist in unserem kapitalistisch-neoliberalen System mehr als wahrscheinlich, dass die Arbeitskraft von KI im Wesentlichen nur dem Prozent der Superreichen zugutekommt und dafür Menschen massenhaft in prekäre Lebensbedingungen stürzen wird. Aber wer kauft dann die von Robotern produzierten Waren? Wie bleiben die Reichen reich ohne funktionierende Marktwirtschaft und Absatzmärkte? Der Markt würde nichts mehr regeln und das ewige Wachstum des Kapitalismus hätte sich selbst gefressen. Das ist sicher nicht die Zukunft, in der wir leben wollen und können. Angesichts der nuklearen Bedrohung im Kalten Krieg hat die Menschheit schon mal zerstörerische Szenarien nicht bis zum Ende durchgespielt – vielleicht gelingt uns das auch diesmal und wir können die immer dramatischer werdende Akkumulation von Kapital umkehren.

Zum Thema KI und Zukunft ist nur eines wirklich klar: Es wird immer Geschichten über fühlende oder rebellische KI geben, denn die Science Fiction dreht sich um Menschen und ihre Beziehungen, Hoffnungen und Ängste im Wechselspiel mit Technologie, und ist keine Blaupause für maschinengesteuerte Zukünfte.

Ausgezeichnet mit dem

**KURD LASSWITZ**

**PREIS 2024**



Hardcover | 272 Seiten | 32 €  
ISBN 978-3-949452-96-3



Hardcover | 484 Seiten | 25 €  
ISBN 978-3-949452-93-2

Gibt's überall, wo es Bücher gibt,  
und direkt bei uns:  
**shop.hirnkost.de**



**HIRNKOST**  
Engagierte Literatur

Dietmar Dath

# ES SCHREIBT UND SPRICHT. HAT'S EIN GESICHT?

Über Sprache, Text und Denken als Dreieck in  
der Science Fiction

## Eins: Bedienung und Bewusstsein

Der im April 2024 veröffentlichte SF-Roman *Welcome to Boy.net* von Lyda Morehouse ist alles andere als eine schwerfällige oder weit-schweifige Abhandlung über die Bedeutung der Auslegung von Texten für Wissenschaft, Technik und Literatur. Im munteren Fortgang der erzählten Handlung (es geht um zwei Weltraum-Kopfgeldjägerinnen, die einander lieben und alles zerlegen, was sich dieser Liebe in den Weg stellt) übersieht man deshalb leicht, dass Morehouse in dem Buch einen ehrgeizig weitreichenden Vorschlag dazu untergebracht hat, die zur Zeit des Erscheinens des Romans an vielen gesellschaftlichen Fronten sehr lebhaft geführte Diskussion über künstliche Intelligenz von einem entscheidenden Punkt her neu aufzuziehen.

Die Stelle, an der dieses Thema explizit wird, beginnt mit einer literarischen Anspielung: Eine Maschine namens Margie, die in einem Restaurant die Kundschaft bedient, äußert ihr Missfallen darüber, dass ihr »Gehirn von der Größe eines Planeten« verschwendet sei an diesen Kellnerjob. Hawk, Co-Heldin des Romans, erkennt das indirekte Zitat, das, wie sie weiß, aus einem »English humor classic about a galactic hitchhiker« stammt, nämlich aus dem Hauptwerk des SF-Satirikers Douglas Adams, der aus einem Hörspielprojekt entsprungenen Pentalogie *THE HITCHHIKER'S GUIDE TO THE GALAXY*, deren fünf Bände zwischen 1979 und 1992 erschienen.

In Adams' lustiger Welt gibt es den Roboter Marvin, den seine chronische Unterforderung in depressive Monologe treibt, in denen unter anderem der Vergleich des avancierten künstlichen Gehirns mit einem Planeten vorkommt, nämlich mit der Gesamtgedankenfülle der Wesen, die so einen Planeten bevölkern könnten.

Margies Marvin-Mini-Rollenspiel weckt Hawks Argwohn: Soll das etwa heißen, die Serviererin behauptet, sie besäße echte künstliche Intelligenz, also ein Bewusstsein von sich selbst? Das kann nicht sein, denkt Hawk, denn die Technik und die Wissenschaft ihrer Epoche, in der die irdische Politik sich auf einen Großteil des Sonnensystems ausgeweitet hat, haben bislang nichts Intelligenteres zuwege gebracht als ein paar quasi-autonome Puppen, deren Beliebteste sexuellen Gebrauchswert bieten (sie heißen Ken, Barbie oder Bowie).

»They faked emotions and seemed real to most people«, meint Hawk abfällig, biologischen Leuten werden von synthetischen also Emotionen nur vorgespielt, per Kommunikation und sonstigem Verhalten, und das reicht den Leuten, die diese Geschöpfe benutzen.

### **Zwei: Alan Turings tückische Idee und deren Überwindung**

Die etwas selbstgefällige Haltung, die Hawk da einnimmt, ist unterfüttert von technikpolitisch-zeitgeschichtlichen Vorannahmen, zu denen Morehouse uns mitteilt, dass eine dazu befugte Körperschaft sie in der Welt ihres Romans als geltendes Recht kodifiziert hat. Eine »Lunar Conference«, also eine große Tagung auf dem Mond, von deren Beschlüssen offenbar alle, die es angeht, alles wissen, hat nämlich eine Definition von »true intelligence«, also »echter Intelligenz« aufgerichtet, die den Zweck hat, das berühmte Gedankenexperiment des sogenannten »Imitationsspiels« zu überwinden, mit dem heute, also in der realen Ära, in der Lyda Morehouses Texte geschrieben, gedruckt und online gestellt werden, noch weithin gearbeitet wird, wo immer Menschen zu verstehen versuchen, was die Sprach-, Symbol- und Kalkülnutzung von beispielsweise Deep-Learning-Algorithmen von menschlichen Verstandes- und Kommunikationsgebrauchsweisen einerseits unterscheidet und sie andererseits mit den Hervorbringungen unserer Gehirne vergleichbar macht oder anderweitig verbindet. Jenes Imitationsspiel hat der Erfinder der abstrakten Begrifflichkeit universeller Rechenmaschinen, der Engländer Alan Turing, 1950 ersonnen. Er wollte die Frage »Können Maschinen denken?« so weit präzisieren, dass sie für konkrete Maschinen beantwortbar wird.

Der Dreh geht so: Wenn im Laufe eines Kommunikationsgeschehens zwischen einem Menschen A und einem abgeschirmten Gegenüber B, das weder eindeutig als Mensch noch als Maschine zu

erkennen ist, der Mensch A nicht herausfinden kann, ob das Gegenüber B Mensch oder Maschine ist, dann darf man von B, sofern es sich dabei tatsächlich um eine Maschine handelt, fortan behaupten, dieses Wesen könne denken.

Die heikelste Implikation des Spiels ist der Satz: Wer so gut lügt, wie wir das können, der denkt auch sonst wie unsereins.

Turings Einfall ist lupenreine SF – ach was, von wegen Lupe, er ist rastertunnelmikroskopreine SF: kompakter, straffer, dichter kann man das Problem nicht erzählbar machen. Kolleginnen wie Kollegen aus dem Genre haben die Idee nur allzugern aufgegriffen, teils allerdings, um sie von innen her zu zerstören; man sollte hier an Philip K. Dicks Androidenszenarien denken oder an Ian McDonalds ebenso schlaue wie böse Feststellung, dass jede Maschine, die klug genug ist, den Turing-Test zu bestehen, damit auch klug genug sein dürfte, ihn nicht zu bestehen und die Menschen lieber über ihre Intelligenz zu täuschen, was den Test im Grunde wertlos macht. Hierher sortiert man also auch Lyda Morehouses tiefe, in *Welcome to Boy.net* verarbeitete Einsicht, dass es unerfreulicherweise zum evolutionär-historischen Gepäck unserer Gattung gehört, ihre Umgebung projektiv zu beleben (Animismus!) und sogar zu vermenschlichen: Kaum malt jemand ein Menschengesicht auf einen Fetisch, schon meinen wir, einen Atemhauch zu spüren, der von dem Ding ausgeht. Es gibt arme Seelen, die mit Sexpuppen zusammenleben oder ein mit menschlichen Zügen versehenes Kissen für eine anders nicht aufzutreibende Gefährtin oder einen anders nicht erlangbaren Gefährten halten wollen. Wir sind, sagt Morehouse, zu schwach für die Disziplin, die nötig ist, der Versuchung zur Selbsttäuschung zu widerstehen, die der Turing-Test jeder Person A nahelegt. Deshalb ist die fiktive Mondkonferenz bei der Frage »Können Maschinen denken?« zu einem anderen Kriterium als dem von Turing entwickelten gelangt: »the only thing they would consider a true intelligence was a program that could resist revision – a program that not only said that it wanted to exist, but which could actively fight for itself. If you could get down into the code, change a few specs and end up with a different personality, then that personality was never ›reak. The Lunar Conference codified the idea that the only true human act was *resistance*.«

Wenn du in den Code, in den Programmtext, eindringen kannst und ein paar Spezifika ändern (bei den derzeit breit diskutierten Chatbots

wären das etwa die Wahrscheinlichkeits-Leitgrößen beim Generieren von Ausgaben auf Prompts hin, die man in dem Begriff »Temperatur« bündelt, oder die Gewichte, also die Stärken der Verbindungen zwischen den Rechenzellen der künstlichen neuronalen Netze), und wenn der Code sich dann nach diesen Eingriffen anders benimmt als vorher, dann »fühlt« er sich nicht plötzlich anders, weil er sich nämlich nie irgendwie »gefühl« hat, es war alles Eingabe-Ausgabe-Probabilistik, niemand zu Hause. So proklamiert das die Mondkonferenz, woraus für Hawk dann der Dünkel gegenüber der Kellnerin Margie folgt, auch wenn diese Maschine einen pffiffigen Witz unter Rückgriff auf eine Datenbank reißt, in der irgendetwas von Douglas Adams und seiner Schöpfung Marvin zu finden ist. Nein, findet Hawk, in archaisierendem Englisch mit gravitätischer Grammatik, »one sarcastic and seemingly response did not an artificial intelligence make.« Soll sagen: eine einzelne smarte Sprachhandlung macht so wenig eine echte Künstliche Intelligenz wie eine Schwalbe einen Sommer.

### **Drei: Schreiben, Lesen, Wissen, Wollen**

Computerprogramme sind Texte.

Code wird geschrieben und gelesen (in zunehmendem Maß freilich ist beides Metapher, denn es sind in letzter Zeit an diesen Vorgängen keineswegs immer Menschen beteiligt), Code wird übersetzt (in andere Sprachen für Rechner, in Ausführungsbestimmungen und so fort), Code wird manchmal gedeutet. Die Eingriffe, die bei Morehouse klären sollen, ob Code denkt, werden in unseren Tagen, in der Realität der Mitte der Zwanzigerjahre des 21. Jahrhunderts christlich-abendländischer Zeitrechnung, gerade zu mancherlei diagnostischen Zwecken entwickelt und rasch weitergetrieben, weil wir sonst nicht nur im Unklaren darüber bleiben, ob die von uns geschaffenen Systeme irgendetwas verstehen können, sondern bald selbst kaum noch verstehen, was diese Systeme tun. Da entsteht eine neue Forschungsdisziplin namens »Explainable Artificial Intelligence«, erklärbare künstliche Intelligenz, auch als »erklärbares Maschinenlernen« geläufig, abgekürzt XAI.

Dieser blutjunge Wissenschaftszweig befasst sich etwa mit den Wegen, auf denen eine Eingabe in den hintereinandergeschalteten Schichten von Rechenzellen, aus denen die augenblicklich erfolgreichsten KI-Vorrichtungen, die Deep-Learning-Systeme, bestehen,