



BÜRO FÜR TECHNIKFOLGEN-ABSCHÄTZUNG
BEIM DEUTSCHEN BUNDESTAG

Christoph Kehl
Christopher Coenen

Technologien und Visionen der Mensch-Maschine- Entgrenzung

Sachstandsbericht zum TA-Projekt
»Mensch-Maschine-Entgrenzungen:
zwischen künstlicher Intelligenz und
Human Enhancement«

Juli 2016
Arbeitsbericht Nr. 167





**Technologien und
Visionen der
Mensch-Maschine-
Entgrenzung**



Das Büro für Technikfolgen-Abschätzung beim Deutschen Bundestag (TAB) berät das Parlament und seine Ausschüsse seit 1990 in Fragen des technischen und gesellschaftlichen Wandels. Das TAB ist eine organisatorische Einheit des Instituts für Technikfolgenabschätzung und Systemanalyse (ITAS) im Karlsruher Institut für Technologie (KIT). Zur Erfüllung seiner Aufgaben kooperiert es seit September 2013 mit dem Helmholtz-Zentrum für Umweltforschung GmbH – UFZ, dem IZT – Institut für Zukunftsstudien und Technologiebewertung gGmbH sowie der VDI/VDE Innovation + Technik GmbH.



Christoph Kehl
Christopher Coenen

Technologien und Visionen der Mensch-Maschine- Entgrenzung

Sachstandsbericht zum TA-Projekt
»Mensch-Maschine-Entgrenzungen:
zwischen künstlicher Intelligenz und
Human Enhancement«

TAB-Arbeitsbericht Nr. 167



Büro für Technikfolgen-Abschätzung
beim Deutschen Bundestag (TAB)
Neue Schönhauser Straße 10
10178 Berlin

Fon: +49 30 28491-0

Fax: +49 30 28491-119

buero@tab-beim-bundestag.de

www.tab-beim-bundestag.de

2016

Umschlagbild

aliencaat © 123RF.com

Gedruckt auf Circle offset Premium white.

ISSN-Print 2364-2599

ISSN-Internet 2364-2602



Inhalt

Zusammenfassung	7
<hr/>	
I. Einleitung	19
<hr/>	
II. Zukunftsvisionen	25
1. Zur Analyse des visionären Diskurses	26
2. Kulturelle Wurzeln und Ideentraditionen	29
2.1 Künstliche Menschen, Automaten und Roboter	29
2.2 Transhumanismus, Cyborgs und Human Enhancement	33
3. Aktuelle visionäre Diskurse und ihre Akteure	37
3.1 Themensetzung durch futuristische Netzwerke	38
3.2 Der aktuelle Diskurs zu Enhancement und Cyborgs	43
3.3 Der aktuelle Diskurs zu einer Superintelligenz	47
4. Fazit	50
<hr/>	
III. Stand der Technik	55
1. Neurotechnologien	55
1.1 Wissenschaftliche Grundlagen: neuroelektrische Schnittstellen	56
1.2 Neurotechnologische Anwendungen: Stand und Entwicklungsperspektiven	67
1.3 Forschungslandschaft in Deutschland	93
1.4 Fazit und Ausblick	95
2. Autonome Robotik	99
2.1 Wissenschaftliche Grundlagen: künstliche Intelligenz und maschinelles Lernen	102
2.2 Wo steht die autonome Robotik?	110
2.3 Anwendungsfelder und Beispiele	120
2.4 Forschungslandschaft in Deutschland	134
2.5 Fazit	137
<hr/>	
IV. Schlussfolgerungen und Ausblick	141



Literatur	153
In Auftrag gegebene Gutachten	153
Weitere Literatur	153
<hr/>	
Anhang	167
1. Abbildungsverzeichnis	167
2. Tabellenverzeichnis	167

Zusammenfassung

Als Mensch-Maschine-Entgrenzung wird im Folgenden das Phänomen bezeichnet, dass sich die ursprünglich klaren Grenzen zwischen Menschen und den von ihnen geschaffenen technischen Arbeitsmitteln, den Maschinen, zunehmend aufzulösen beginnen. Dazu beigetragen haben die großen Erkenntnisfortschritte, die in verschiedenen, nur lose miteinander verknüpften, aber immer stärker konvergierenden wissenschaftlichen Bereichen in den letzten Jahren und Jahrzehnten gemacht worden sind. So ermöglichen die Neurowissenschaften Einsichten und Eingriffe in das Gehirn, die noch vor Kurzem undenkbar schienen, während das Forschungsfeld der künstlichen Intelligenz (KI) zunehmend erfolgreich an der Nachbildung lernfähigen und intelligenten Verhaltens arbeitet. Zusammen mit Fortschritten in den Informations- und Neurotechnologien eröffnen sich dadurch zukunftssträchtige Anwendungsperspektiven, die das Potenzial haben, das Mensch-Technik-Verhältnis ganz neu zu definieren. Eine Schlüsselrolle spielen dabei immer autonom agierende Maschinen, die als kollaborative Roboter über zunehmend menschliche Züge und Fertigkeiten verfügen oder als (extern oder intern getragene) Neuroprothesen mit dem menschlichen Körper verschmelzen. Auch wenn die technologische Entwicklung erst in den Anfängen steckt, begleitet und angetrieben wird sie gleichwohl von weitreichenden Zukunftsvorstellungen, welche bei aller Unschärfe und Widersprüchlichkeit die große gesellschaftliche Tragweite dieser Technisierungsprozesse deutlich vor Augen führen. Der vorliegende Sachstandsbericht hat u. a. zum Ziel, auf Basis einer fundierten Analyse des gegenwärtigen Standes von Forschung und Technik dem Realitätsgehalt dieser Visionen auf den Grund zu gehen.

Zukunftsvisionen

Die öffentlichen Debatten zu Technologien der Mensch-Maschine-Entgrenzung werden in hohem Maße durch weitreichende Zukunftsvisionen beeinflusst, die seit Anfang der 2000er Jahre vielfältige Beachtung finden. Diese Visionen sind kulturell sehr wirkmächtig und haben ganze Genres, wie z. B. die Science-Fiction (in der Literatur, im Film und Spielbereich), maßgeblich geprägt. Vor diesem Hintergrund können die gegenwärtigen Diskussionen über Human Enhancement (HE) und über KI als zwei Paradebeispiele für die in den letzten Jahren zunehmend vertretene Ansicht gelten, dass bei der Ab- und Einschätzung neuer oder in der Entstehung befindlicher (emergierender) Technologien zumeist auch eine systematische Untersuchung von Zukunftsvisionen angezeigt ist.

Betrachtet man allein die naturwissenschaftlich-technischen Grundlagen und Anknüpfungspunkte der visionären Diskussionen, liegt es nicht unbedingt nahe, Zukunftsvisionen zum neurotechnologischen Enhancement sowie zur KI und Robotik gemeinsam zu betrachten. Wie der Bericht zeigt, gibt es aber nicht nur auf der Anwendungs-, sondern auch auf der diskursiven Ebene etliche Berührungspunkte. So spielt der sogenannte Transhumanismus, dem es um eine radikale Veränderung oder gar Überwindung der *Conditio humana* geht, als ein übergreifendes Deutungsmuster eine zunehmend zentrale Rolle. Auch auf der Akteursebene sind hinsichtlich der Visionsproduzenten und -verbreiter und ihrer Unterstützer erhebliche Überlappungen der beiden Felder festzustellen.

Im Wesentlichen haben sich die visionären Diskussionen über die Zukunftsaussichten »starker« KI und über neurotechnologiebasierte Mensch-Maschine-Entgrenzungen nach einem Muster entwickelt, das aus älteren zukunftsvisionären Technikdebatten – z. B. über Nanotechnologie – bekannt ist. In mehr oder weniger eng miteinander verbundenen, relativ kleinen Zirkeln namhafter Naturwissenschaftler, Ingenieure, Autoren und Industrieller werden z. T. sehr weitreichende Zukunftserwartungen diskutiert. Diese Erwartungen sind jedoch – wie die Analyse im Bericht zeigt – weitgehend deckungsgleich mit Spekulationen, die bereits vor vielen Jahrzehnten angestellt wurden. Im Fall der aktuellen Debatte über »Superintelligenz« (bzw. eine sehr leistungsfähige KI) wurden im Wesentlichen Motive aus den 1960er Jahren aufgegriffen – etwa die Idee einer selbstbewussten und sich selbstverstärkenden KI (Intelligenzexplosions- oder heute zumeist Singularitätshypothese) des Mathematikers Irving J. Good. Auch die aktuellen Diskussionen über ein Enhancement mit neurotechnologischen Mitteln haben eine lange Vorgeschichte. Ein grundlegendes Element ist die Interpretation des Menschen als ein Mängelwesen, vor deren Hintergrund dann z. B. der Physiker John D. Bernal, ein Pionier der modernen Lebenswissenschaften und des transhumanistischen Denkens, im Jahr 1929 die Vision einer massiven, auch neurotechnologisch ermöglichten Cyborgisierung des menschlichen Körpers und einer letztendlichen Überwindung seiner natürlichen Beschränkungen entwarf. Neu – z. B. im Vergleich zum Diskurs über Nanotechnologie – ist jedoch zum einen, dass führende Figuren der Computer- und Internetindustrie (z. B. der PayPal-Gründer Elon Musk) nun selbst die Debatten inhaltlich mitprägen sowie die Forschung zu technikvisionären Themen und transhumanistische Aktivisten noch stärker finanziell fördern. Zum anderen hat sich der Diskurs weit über die (vor allem in Großbritannien und den USA) seit längerer Zeit aktiven technikvisionären Zirkel hinaus ausgebreitet und auch in Deutschland zunehmend Beachtung gefunden, nicht zuletzt in den Massenmedien.

Der Umstand, dass ein relativ kleiner Kreis von Akteuren die zukunftsvisionären Diskussionen wesentlich prägt, sollte einerseits nicht zum Anlass genommen werden, die Diskussionen über Mensch-Maschine-Entgrenzungen

zu ignorieren. Andererseits sollten aber der erhebliche intellektuelle Einfluss oder die ökonomische Macht vieler Beteiligten nicht dazu verleiten, ohne nähere Prüfung eine besondere Dringlichkeit oder Bedeutung der verhandelten Themen vorauszusetzen. Letztlich haben wir es hier fast durchgängig mit z. T. schon sehr alten Zukunftsvisionen zu tun, deren tatsächliche Basis angesichts des aktuellen Forschungs- und Entwicklungsstandes aber immer wieder weiterer Klärung bedarf. Neben Analysen, in denen die Zukunftsvisionen an der Realität gemessen werden, erscheinen auch Anstrengungen angezeigt, die inhärenten Mechanismen und Dynamiken technikfuturistisch geprägter Diskussionen besser zu verstehen. Dies betrifft vor allem den Aspekt der Themensetzung durch futuristische Netzwerke und die Logik der massenmedialen Berichterstattung. Der Blick auf die aktuelle Debatte über »starke« KI etwa zeigt, dass zwar viele Prominente vor der Entstehung einer »starken« KI mit Ichbewusstsein warnen, womit sie gleichwohl dazu beitragen, bestimmte Technikvisionen zu popularisieren und Aufmerksamkeit für noch nichtexistierende Technologien zu schaffen.

Stand und Perspektiven der Neurotechnologien

Die Reiz- und Signalverarbeitung im menschlichen Nervensystem erfolgt primär in Form elektrischer Erregung. Neurotechnologien machen sich dieses neurophysiologische Wirkprinzip zunutze, um maschinelle Systeme über neuroelektrische Schnittstellen mit dem menschlichen Nervensystem zu koppeln. Im Falle *stimulierender Schnittstellen* werden elektrische Signale dabei vom Apparat zum Nervensystem, im Falle *ableitender Schnittstellen* vom Nervensystem zum Apparat übertragen. Außerdem ist zwischen invasiven und nichtinvasiven Verfahren zu unterscheiden, je nachdem, ob sich die Schnittstelle im Inneren des Körpers befindet, was einen operativen Eingriff erforderlich macht, oder auf der Körperoberfläche, was zwar in der Anwendung unkomplizierter ist, dafür jedoch unspezifischere Signale zur Folge hat. Durch diese verschiedenen Typen an neuroelektrischen Schnittstellen ergibt sich eine große Bandbreite an möglichen Anwendungen, hauptsächlich im klinischen Bereich, die sich in unterschiedlichen Phasen der Entwicklung bzw. Anwendungsreife befinden.

Ableitende Systeme unterscheiden sich danach, wo neuronale Aktivität erfasst wird (im peripheren oder zentralen Nervensystem) und auf welche Art (invasiv vs. nichtinvasiv), sowie natürlich auch danach, welche Geräte letztendlich gesteuert werden sollen (Prothese, Rollstuhl, Computer etc.). Die Steuerung von Greifprothesen ist einer der zentralen und angesichts der Vielzahl der Betroffenen auch kommerziell aussichtsreichsten Anwendungsbereiche. Seit über 30 Jahren behilft man sich hier mit myoelektrischen Prothesen, bei denen Elektroden die verbliebene Muskelaktivität im Stumpf erfassen und damit batterie-

betriebene Motoren in Gang setzen. Besonders bei der technischen Nachbildung der Hand mit ihren verschiedenen Bewegungsmöglichkeiten stößt diese Methode aber aufgrund schwacher Signale schnell an Grenzen. Abhilfe verspricht eine Steuerung direkt durch die Hirnaktivität, also quasi durch Gedankenkraft, ohne den Umweg über das periphere Nervensystem. Dieser Weg wird seit etwa 20 Jahren intensiv erforscht. In dem vorliegenden Bericht werden diesbezügliche Anwendungen, die nichtinvasive Hirnschnittstellen umfassen, als Brain-Computer-Interfaces (BCI) bezeichnet, solche mit invasiven Schnittstellen als Brain-Machine-Interfaces (BMI):

- › *BCI-Systeme* nutzen in der Regel die neuroelektrische Gehirnaktivität, die über Elektroenzephalogramm (EEG) an der Kopfoberfläche erfasst und in maschinelle Steuerkommandos übersetzt wird. Sie sind bislang vor allem als computerbasiertes Kommunikationsmittel (z. B. Buchstabierhilfe bzw. Surfen im Internet) für schwer gelähmte Patienten mit einigem Erfolg erprobt worden. Andere Anwendungszwecke – insbesondere die Steuerung externer Geräte wie Rollstühle, Hilfsroboter, Exoskelette, Drohnen etc. – werden zwar seit vielen Jahren intensiv erforscht, haben es aber noch nicht bis zur Anwendungsreife gebracht. Schnelle Erfolge sind auch in nächster Zeit nicht zu erwarten. Denn die Gewinnung von Oberflächenpotenzialen auf der Kopfhaut von gesunden wie kranken Nutzern ist zwar unproblematisch und weitgehend risikolos, leidet gleichzeitig aber an geringer räumlicher Auflösung sowie an verrauschten Signalen mit geringer Bandbreite. Die Nichtinvasivität erweist sich folglich zugleich als größter Pluspunkt wie auch als wesentlicher Nachteil derartiger Anwendungen.
- › *BMI-Systeme*: Im Unterschied zu den räumlich schlecht aufgelösten und oft verzerrten EEG-Ableitungen lassen sich mittels implantierter Mikroelektroden prinzipiell sehr hohe und spezifische Datenraten aus dem Gehirn gewinnen. Aus diesem Grund eignen sich BMI-Anwendungen besser dazu, auch komplexere Gerätschaften wie etwa avancierte Gliedmaßenprothesen zu kontrollieren. Allerdings sind Prothesen, die direkt mit dem Gehirn verkabelt sind, noch nicht aus dem Stadium der Grundlagenforschung herausgekommen und wurden bislang vor allem im Tierversuch und nur in Einzelfällen beim Menschen prototypisch realisiert. Einer verbreiteten Anwendung steht derzeit neben den operativen Risiken vor allem die noch sehr eingeschränkte Langzeitstabilität der implantierten Sonden im Weg, die aufgrund der Fremdkörperreaktion des Körpers (Einkapselung) in der Regel nach wenigen Wochen ihre Funktion verlieren.

Während sich die ableitenden Systeme noch in einem sehr frühen Entwicklungsstadium befinden, was vor allem mit den komplexen Anforderungen an die Datenerfassung und -analyse zusammenhängt, gehören die vom Aufbau her

etwas einfacheren *stimulierenden Systeme* teilweise bereits seit Längerem zum klinischen Alltag. Auch hier ist zwischen invasiven und nichtinvasiven Verfahren zu unterscheiden, wobei vor allem invasive Stimulationsanwendungen – etwa das Cochlea-Implantat zur Wiederherstellung des Gehörs oder die tiefe Hirnstimulation (Deep Brain Stimulation [DBS]) zur symptomatischen Therapie neuromotorischer Erkrankungen wie Parkinson – zum festen Standardrepertoire der Medizin gehören. Der Vorteil von invasiv stimulierenden Schnittstellen, die neben den bereits erwähnten noch viele andere therapeutische Anwendungen haben (z. B. als Rückenmark-, Muskel-, Blasen- oder Enddarmstimulatoren), liegt darin, dass im Unterschied zu den derzeit erforschten nichtinvasiven Stimulationsverfahren, wie der transkraniellen Magnet- oder der Elektrostimulation, einzelne Muskeln, Nerven oder Hirnbereiche wesentlich spezifischer gereizt werden können.

Speziell die weite Verbreitung sensorischer Neuroprothesen und der DBS-Therapie, die beide auf das zentrale Nervensystem einwirken, wirft ethische Fragen auf und befeuert Cyborgvisionen, die eine baldige Verschmelzung von Mensch und Technik zu einer Art kybernetischem Mischwesen prognostizieren:

- > Bei der *tiefen Hirnstimulation*, die außer bei Parkinson inzwischen u. a. auch bei schweren Fällen von Zwangsstörungen und Depressionen zum Einsatz kommt, werden dem Patienten winzige Elektroden tief in das Gehirn implantiert und dort verankert, der Stromimpuls erfolgt über einen Impulsgeber, der unterhalb des Schlüsselbeins unter die Haut eingepflanzt wird. Der therapeutische Nutzen der DBS ist bei den erwähnten Indikationen anerkannt, dennoch kommt sie nur als ultima ratio zur Anwendung. Dies hängt zum einen mit der Schwere des Eingriffs, zum anderen aber auch mit grundsätzlichen ethischen Vorbehalten zusammen, da bei Patienten teilweise gravierende psychische Nebenwirkungen bis hin zu Persönlichkeitsveränderungen beobachtet wurden.
- > *Sensorische Prothesen*, wie das weitgehend ausgereifte Cochlea-Implantat oder das seit Kurzem erhältliche Retina-Implantat (zur Wiederherstellung des Sehsinns), setzen an peripheren Sinnesnerven an, um durch Reizung derselben die verlorengegangene Wahrnehmungsfähigkeit zumindest rudimentär zu kompensieren. Dabei verschmelzen menschliches Gehirn und technisches Hilfsmittel zu einer Funktionseinheit, wobei auch gewisse Sinneserweiterungen möglich (z. B. ein überdurchschnittlich gutes Hören in sehr lauten Umgebungen) oder für die Zukunft denkbar sind (z. B. das Hören von Ultraschall).

Zwar bieten die bestehenden Neurotechnologien also bereits jetzt Optionen, die menschliche Sinneswahrnehmung über das natürlich gegebene Maß hinaus zu erweitern. Es wäre jedoch sehr fragwürdig, solche technischen Sinneserweite-

rungen als eine wünschenswerte Verbesserung der menschlichen Natur darzustellen, weil der erreichbare Nutzen gegenüber den damit verbundenen gesundheitlichen Risiken gering erscheint und daher keine breitere gesellschaftliche Anwendung bei Nichtkranken rechtfertigen und auch in näherer Zukunft nicht erwarten lässt.

Im Rahmen der invasiven Interaktion mit dem Nervensystem ist in jüngster Zeit eine neue Anwendungsdimension am Horizont aufgetaucht, die mit der Entwicklung *bidirektionaler Schnittstellen* einhergeht. Ziel ist eine Integration ableitender und stimulierender Verfahren. Konkret stehen zwei Anwendungsgebiete im Vordergrund:

- › »fühlende« Handprothesen, die eine sensorische Rückmeldung geben und damit ein natürlicheres Greifen ermöglichen sollen;
- › »intelligente« Implantate (auch Elektrozeutika genannt), die Hirnsignale autonom überwachen und nur bei Vorliegen von Auffälligkeiten, die etwa auf einen bevorstehenden epileptischen Anfall hindeuten, stimulatorisch aktiv werden.

Mit der neurotechnologischen Integration von Sensorik und Aktorik wäre eine neue Dimension der Technisierung des Menschen erreicht. Bis derartige invasive Systeme klinisch nutzbar sein werden, dürfte es noch ein langer Weg sein. Wie die Geschichte der Neurotechnologien zeigt, besteht in diesem Feld ein generelles Ungleichgewicht zwischen den intensiven, oft jahrzehntelangen Forschungsbemühungen und den eher spärlichen klinischen Anwendungserfolgen. Verantwortlich dafür sind die hohen Entwicklungshürden, die durch die relativ strengen sicherheitstechnischen Anforderungen an derartige Medizinprodukte noch verschärft werden. Hinzu kommt, dass der Markt für viele neurotechnologische Anwendungen äußerst begrenzt ist, weil die Anwendungen zumeist nur auf sehr spezifische Indikationen zugeschnitten sind und damit letztlich auch nur für überschaubare Patientengruppen infrage kommen.

Stand und Perspektiven der autonomen Robotik

Roboter – gemeinhin als Manipulationsmaschinen definiert, die über einen gewissen Grad an Autonomie verfügen – stehen synonym für die zunehmende Automatisierung der Gesellschaft, die im industriellen Bereich mit dem Aufkommen der ersten Industrieroboter in den 1970er Jahren eingesetzt hat. Hierbei handelt es sich in der Regel um große und schwere Maschinen, die aus Sicherheitsgründen strikt vom Menschen getrennt werden und weitgehend repetitive Tätigkeiten in hochgradig strukturierten Umgebungen zu vollbringen haben – Roboter dieser Art sind in der produzierenden Industrie heute massenhaft verbreitet. In den letzten 10 Jahren haben Forschung und Entwicklung auf dem

Gebiet der Robotik jedoch deutliche Fortschritte gemacht, sodass eine neue Generation von Robotern am Horizont auftaucht, die sogenannten *autonomen Roboter*, die als Service- oder Assistenzroboter zunehmend in alltägliche Bereiche vordringen und dabei immer enger mit Menschen interagieren. Dieser neue Robotertyp verfügt über einen höheren Grad an Selbstständigkeit und ist in der Regel deutlich leichter und flexibler konstruiert als klassische Industrieroboter – er stellt mithin ganz neue Herausforderungen an Forschung und Entwicklung.

Wesentliche Wissensgrundlagen für die Konstruktion autonomer Systeme sind in den beiden eng verwandten Forschungszweigen KI sowie des maschinellen Lernens (ML) verankert. Insbesondere die KI-Forschung, deren Wurzeln bis in die 1940er Jahre zurückgehen, hat in ihrer wechselvollen Geschichte die Entwicklung robotischer Systeme maßgeblich geprägt. Während lange Zeit die Vorstellung einer »starken« KI leitend war, also die künstliche Nachbildung allgemeiner menschlicher Intelligenz inkl. Bewusstsein, hat diese Idee nach etlichen Rückschlägen innerhalb der Wissenschaft inzwischen an Anziehungskraft verloren. Stattdessen stehen heute spezifische Anwendungsprobleme im Vordergrund, die man mithilfe künstlicher Systeme zu lösen versucht – ohne dabei im umfassenden Sinne intelligente Maschinen schaffen zu wollen (sogenannte »schwache« KI). In den Bereich der »schwachen« KI fallen etwa die Forschungen zum ML, ein Gebiet, in dem in den letzten Jahren auf algorithmischer Ebene beeindruckende Anwendungserfolge erzielt werden konnten.

Autonome Roboter lassen sich als Perception-Action-Learning-Systeme charakterisieren, d. h. als physische Agenten, die Wahrnehmung und Verhalten interaktiv und unter Echtzeitbedingungen realisieren. Die Entwicklung derartiger Systeme erfordert eine konsequent interdisziplinäre Herangehensweise, die weit über die skizzierten Grundlagenfächer der KI- und ML-Forschung hinausgeht und vor allem auch technische Disziplinen wie die Mechatronik, die Regelungstechnik oder die Elektrotechnik involviert. Zentrale Teilbereiche von Forschung und Entwicklung im Feld der autonomen Robotik sind der Wahrnehmung, Lern- und Planverfahren (Steuerung) sowie der Systemintegration zugeordnet:

- > *Wahrnehmung*: Grundlage für die Verhaltenssteuerung autonomer Roboter ist die Erstellung eines hinreichend guten Modells der Roboterumgebung und des eigenen Zustands aus sensorischen Informationen. Dafür braucht es erstens vielfältige Sensorik, die interne und externe Parameter erfassen kann, und zweitens Verfahren, die eine semantische Interpretation der wahrgenommenen Objekte ermöglichen (Objektklassifikation). Im Bereich der Sensorik und den Methoden der Dateninterpretation, Mustererkennung und Kartierung (etwa mithilfe von Deep Learning) wurden in den letzten Jahren deutliche Fortschritte erzielt. Dennoch stellt die autonome Wahrnehmung in Echtzeit, vor allem in Bezug auf ein umfassendes Szenenver-

ständnis auf Grundlage einzelner Wahrnehmungsinhalte, immer noch eine große Herausforderung dar. Dies gilt im Besonderen, wenn es darum geht, menschliche Bewegungen zu erkennen und dann die Absicht dieser Bewegungen zu eruieren, was für die Interaktion mit Menschen von großer Bedeutung ist.

- › *Lern- und Planverfahren:* Neben Sensoren und Aktuatoren, die es ermöglichen, die Umwelt zu vermessen und in diese einzugreifen, sind insbesondere Lern- und Adaptionsmechanismen zentral, welche das System befähigen, sich an sich verändernde Umgebungen anzupassen oder komplett neue Verhaltensweisen zu erwerben. Die verfügbaren Lern- und Planungsansätze stoßen in unstrukturierten Umgebungen und bei Bewegungsabläufen, wie sie humanoide Manipulationsroboter (mit ihren oft über mehr als 50 Regelungsdimensionen) zu vollbringen haben, noch definitiv an ihre Grenzen. Erforderlich ist viel manuelle Programmierung und externe Feinsteuerung. Werden heute Roboter in komplexen Domänen eingesetzt, dann agieren diese deshalb in der Regel nur semiautonom: Der Roboter führt einzelne, klar definierte Aktionen autonom aus, die übergeordnete Steuerung sowie Zielfestlegung ist jedoch weiterhin Sache des Menschen.
- › *Systemintegration:* Ein Roboter ist mehr als die Summe seiner Teile. Erst durch die mechanische, elektrische und informationstechnische Verbindung der verschiedenen Systembestandteile entsteht in der Summe sinnvolles Roboterverhalten. Ein wesentlicher Teil der Integration der unterschiedlichen Systemkomponenten zu einem funktionsfähigen Gesamtsystem wird durch die Robotersoftware geleistet. Sie stellt eine der zentralen Hürden bei der Entwicklung autonomer Roboter dar, da vorhandene Softwarelösungen oft nur auf einen einzigen Roboter und einige spezialisierte Anwendungen zugeschnitten sind. Der dafür erforderliche Aufwand ist – nicht zuletzt wegen fehlender Standards im Bereich von Hard- und Software – erheblich und erfordert interdisziplinäre Kompetenzen, über die nur wenige Softwarespezialisten verfügen.

Die autonome Robotik hat in den letzten Jahren zweifelsohne erhebliche Fortschritte gemacht. Angesichts des demografischen Wandels, der die meisten westlichen Staaten erfasst hat, ist eine verbreitete Hoffnung für die Zukunft, dass autonome technische Helfer die dadurch zu erwartenden Produktivitätsverluste ausgleichen können. Das erwartete Marktpotenzial für autonome Roboter ist demnach enorm, und es erstaunt deshalb nicht, dass sowohl private als auch öffentliche Akteure – zu nennen sind insbesondere die US-amerikanische Defense Advanced Research Projects Agency (DARPA) und die großen IT-Konzerne Google, Apple und Amazon – bereit sind, viel Geld zu investieren. Die Anwendungsmöglichkeiten für autonome Roboter sind äußerst vielfältig – das Spektrum reicht vom Katastrophen- oder Kriegseinsatz über die Assistenz

in Pflegeeinrichtungen bis hin zum sozialen Gefährten für den Alltag –, die technologischen Anforderungen im Einzelnen so unterschiedlich wie die Einsatzfelder selber. Um ein genaueres Bild der heterogenen FuE-Dynamik zu bekommen, werden im Bericht exemplarische Anwendungen aus den einzelnen Einsatzbereichen vorgestellt, ihr Entwicklungsstand sowie die sich dabei stellenden Herausforderungen und Lösungsansätze kursorisch diskutiert.

Die Breite der dargestellten Anwendungsbereiche macht deutlich, dass das Feld das Potenzial hat, viele alltägliche Lebensbereiche grundlegend zu verändern. Auch wenn autonome Roboter heute schon erstaunliche Fähigkeiten besitzen, ist dennoch festzuhalten, dass noch starke Beschränkungen bestehen, die auf längere Sicht Gegenstand intensiver Forschungen bleiben dürften. So erstreckt sich der kommerzielle Aufschwung bislang im Wesentlichen auf mobile Roboter, d.h. Systeme, die sich in robotergerechten Umgebungen auf Rädern fortbewegen und nur sehr spezialisierte Fertigkeiten haben. Die »Königsdisziplin« der autonomen Robotik, nämlich humanoide Systeme, die auch komplexere Manipulationsaufgaben in unstrukturierten Situationen selbstständig durchführen können, befindet sich nach wie vor in einem frühen Entwicklungsstadium – und damit auch Assistenz- und Serviceroboter, die über diese Kompetenzen verfügen müssen, um die in sie gesetzten großen Zukunftshoffnungen erfüllen zu können. Viele Einzelprojekte scheitern mangels langfristiger Finanzierung oder weil sie es nicht schaffen, die benötigten interdisziplinären Fachkompetenzen zu bündeln, die für die Bewältigung eines solchen Vorhabens unabdingbar sind.

Vor diesem Hintergrund ist nicht davon auszugehen, dass sich die gesellschaftlichen Veränderungen abrupt vollziehen werden – eher ist mit einer allmählichen Verbreitung autonomer Roboter im Alltag zu rechnen, womit Raum für politische Weichenstellungen und eine gesellschaftliche Gestaltung der Entwicklung bleiben dürfte. In Bezug auf die autonome Robotik scheint dieser Reflexionsprozess aber auch besonders drängend, und er sollte sich nicht nur auf Sicherheitsfragen beschränken. Denn Ziel dieses FuE-Feldes ist es, Maschinen zu erzeugen, die bisher dem Menschen vorbehaltene Tätigkeiten übernehmen. Diese menschenähnlichen Maschinen transportieren (im Rahmen ihrer technologischen Möglichkeiten) ein bestimmtes Bild des Menschen sowie sozialer Beziehungen und werden somit, je mehr sie in unseren Alltag Einzug halten, auch unser Selbstverständnis und unsere Lebensweise grundlegend herausfordern.

Schlussfolgerungen

Noch entwickeln sich die beiden Technikfelder (Neurotechnologien und autonome Robotik) weitgehend abgegrenzt voneinander, was aufgrund der sehr unterschiedlichen wissenschaftlich-technologischen Grundlagen auch nicht weiter

erstaunlich ist. Aus übergeordneter Sicht ist mit Blick auf ihre Implikationen jedoch hervorzuheben, dass ihre Anwendungsperspektiven die beiden Pole einer übergreifenden Entgrenzungsdynamik markieren: Maschinen werden immer menschenähnlicher (Roboter), gleichzeitig werden menschliche Körper zunehmend technisiert (Cyborg). Auf die Spitze getrieben wird diese Entwicklung durch die sich immer deutlicher abzeichnende Konvergenz von Robotik bzw. KI und Neurotechnologien. So sind Prothesen, welche die Bewegungsabsichten des Trägers aus seinen Muskelsignalen grob »herauslesen«, bereits Standard. Neuerdings wird aber auch intensiv an sogenannten intelligenten Roboterprothesen geforscht, die über einen wesentlich größeren Grad an Autonomie verfügen, um auch bei schwachen Signalen die volle Funktionsfähigkeit von Hand, Arm oder Bein wiederherzustellen. Daneben beginnen sich weitere Anwendungen im Schnittfeld von Robotik und Neurotechnologien herauszukristallisieren, etwa »intelligente« Implantate (Elektrozeutika), robotische Exoskelette oder BCI-gesteuerte Roboter.

Das visionäre Potenzial dieser Entwicklungen steht außer Frage. Bei genauerer Betrachtung zeigt sich jedoch nicht selten, dass sich die populären, häufig transhumanistisch inspirierten Zukunftsvisionen auf Einschätzungen der Entwicklungsdynamik stützen, die bestenfalls hochkontrovers, manchmal auch unklar sind oder gar jeglicher Grundlage entbehren. Derartig überschießende Erwartungshaltungen sind nicht zuletzt ein Symptom der noch sehr unscharfen Konturen der technologischen Felder selbst, die hinsichtlich möglicher Anwendungsperspektiven einen weiten Zukunftshorizont eröffnen und damit allerlei Spekulationen Nahrung bieten. Vor diesem Hintergrund kommt der vorliegende Bericht auf Basis der technologischen Bestandsaufnahme zum Schluss, dass zwischen den realen Anwendungsmöglichkeiten und den teils sehr weitreichenden Visionen ein deutliches Missverhältnis besteht:

- › So bieten auch die klinisch etablierten Neurotechnologien wie die tiefe Hirnstimulation oder das Cochlea-Implantat nur eine gewisse Kompensation für massive krankheitsbedingte Einschränkungen und sind zudem keineswegs risikolos. Die meisten anderen Anwendungen, etwa im Bereich der Gehirn-Maschine-Schnittstellen, befinden sich noch in einem rein experimentellen Stadium, und es wird noch viel Forschungs- und Entwicklungsarbeit erforderlich sein, diese Technologien zur Marktreife zu bringen.
- › Was die KI und Robotik angeht, so scheint die Aussicht auf eine Machtübernahme künstlicher Intelligenzen derzeit vernachlässigbar zu sein angesichts der großen technologischen Herausforderungen (insbesondere der funktionalen Integration von Wahrnehmungs-, Planungs- und Manipulationsfähigkeiten), vor denen die Entwicklung komplexerer Serviceroboter steht. Die bisherigen Erfolge, welche die KI-Forschung unzweifelhaft vorzuweisen hat, beschränken sich auf »lernfähige« Softwareanwendungen –

intelligentes oder autonomes Verhalten im menschlichen Sinne zeigen diese in keiner Weise, sodass völlig unklar bleibt, ob und wie sich eine physisch verkörperte »starke« KI überhaupt realisieren lässt.

Die zu beobachtende Verkürzung der öffentlichen Debatte auf Cyborgs und künstliche Superintelligenzen ist vor allem deshalb wenig hilfreich, weil sie den Blick von den eigentlich drängenden normativen Fragen ablenkt, die mit der weiteren Verbreitung autonomer Systeme bereits jetzt relevant werden. Diese liegen, so ein Befund des vorliegenden Berichts, vor allem in den zunehmenden anthropologischen Grenzerfahrungen, die mit der kategorialen Angleichung von Mensch und Maschine im Zuge der beschriebenen Technisierungsprozesse einhergehen. Eine zentrale Rolle dabei spielt der Begriff der Handlung, der nach gängigen ethisch-rechtlichen Auffassungen eng mit dem Mensch- bzw. Personensein verknüpft ist. So gelten ausschließlich Personen als handlungsfähig, weil nur ihnen die Fähigkeit zugeschrieben wird, sich aus freien Stücken in ihrem Tun Ziele zu setzen und die Mittel auszuwählen, um diese zu erreichen. Maschinen hingegen, als Produkte des Menschen, sind nach herkömmlichem Verständnis rein technische Hilfsmittel, um menschliche Zwecke zu erfüllen. Diese diametrale Abgrenzung von Mensch und Maschine bildet einen wesentlichen Pfeiler unserer normativen Ordnung, etwa in Bezug auf moralische Verantwortungszuschreibungen, die auch rechtlichen Kategorien zugrunde liegen. Der weitere Fortschritt bei Robotern und Neurotechnologien bringt diese Grundordnung immer mehr ins Wanken, insofern nämlich, als Technik im Zuge dessen ihren rein instrumentellen Charakter verliert und als zunehmend selbstständig agierender Akteur ihrerseits tief in das persönlich-individuelle wie auch das gesellschaftlich-kollektive Selbstbewusstsein einzugreifen beginnt. Dass diese Entwicklung vielfältige normative Unsicherheiten aufwirft – etwa im Hinblick auf das leitende Menschenbild, den wünschbaren Grad an gesellschaftlicher Automatisierung (und damit auch Normierung) sowie die Gestaltung des Mensch-Technik-Verhältnisses –, dürfte auf der Hand liegen.

Diese normativen Herausforderungen machen eine frühzeitige Auseinandersetzung mit den gesellschaftlichen Konsequenzen, aber auch Rahmenbedingungen der Entgrenzungsdynamik erforderlich. Wie eine »antizipatorische Governance«, die nicht nur Wissenschaft und Technik, sondern auch Politik, Wirtschaft und potenzielle Nutzer einbeziehen sollte, organisiert werden könnte, ist aber noch eine weitgehend offene Frage. Aus der Tatsache, dass die normativen Implikationen der neuen Entgrenzungstechnologien nicht unabhängig von ihren Anwendungskontexten zu beurteilen sind, lässt sich zumindest folgern, dass eine übergeordnete Debatte zu den Chancen und Risiken dieser Entwicklungen wenig sinnvoll erscheint. Stattdessen ist die gesellschaftliche Auseinandersetzung sinnvollerweise mit einem differenzierten Blick auf die jeweiligen Anwendungsfelder und ihre moralischen Implikationen zu führen, dabei

möglichst unter Mitwirkung relevanter Akteure und Interessenvertreter. Angesichts der anthropologischen Herausforderungen durch die neuen Technologien sollte es dabei letztlich immer auch um die Frage gehen, was wir als genuine Wesenszug des Menschen und seiner Kultur verstehen sowie was davon wir für die Zukunft erhalten und vor maschinellem Zugriff schützen wollen.

Angesichts des demografischen Wandels gilt der Bereich der Pflege und Gesundheit für Deutschland zu Recht als paradigmatisches Anwendungsfeld für Technologien der Mensch-Maschine-Entgrenzung. Denn zukünftig ist mit einer zunehmenden Alterung der Bevölkerung und damit mit einem wachsenden Anteil pflegebedürftiger Menschen an der Gesamtbevölkerung zu rechnen. Der dadurch drohende Pflegenotstand gilt als eine der größten gesellschaftlichen Herausforderungen, die – so wird postuliert – nur mit dem verstärkten Einsatz neuer Technologien zu bewältigen ist. Speziell autonom agierenden Service- bzw. Pflegerobotern sowie der Kombination assistiver Technologien mit nicht-invasiven Neurotechnologien (Steuerung von Exoskeletten, Kommunikationsgeräten etc. mittels Brain-Computer-Interface) wird großes Potenzial zugeschrieben, Pflegekräfte entlasten und Pflegebedürftige im Alltag und bei der Rehabilitation unterstützen zu können – derartige Anwendungen befinden sich in Entwicklung und Erprobung. Wie entsprechende Assistenzsysteme und vor allem auch eine den pflegerischen Herausforderungen angemessene Governance der Technikentwicklung in der Praxis aussehen könnten, sind Fragen, welche in einem anschließenden TA-Projekt vertieft untersucht werden sollen.

Einleitung

I.

Die fortlaufende technische Gestaltung seiner Lebenswelt gilt als ein Wesensmerkmal des Menschen, der aus anthropologischer Sicht deshalb auch gerne als *Homo faber* – als Handwerker – bezeichnet wird. Im Laufe seiner Geschichte schuf der Mensch immer komplexere und wirkmächtigere Artefakte, um in seine Umwelt eingreifen und sie zu seinen Gunsten verändern zu können, wobei mehrere technische Entwicklungssprünge zu verzeichnen waren. Als besonders einschneidend gelten die erste landwirtschaftliche Revolution (ca. 10.000 v. Chr.), in der sich die Techniken zur agrarischen Nahrungsmittelproduktion entwickelten, sowie die industrielle Revolution im 19. Jahrhundert, welche die Dampfmaschine hervorbrachte und mit ihr neue Formen der industriellen Massenproduktion. Diese Technisierungsprozesse haben den menschlichen Alltag tiefgreifend verändert und prägen ihn noch heute. Derzeit befinden wir uns womöglich mitten in einem weiteren fundamentalen, technisch bedingten Umbruchprozess, der als digitale oder informationelle Revolution bezeichnet wird und mit dem das Zeitalter der informationstechnisch vernetzten, teils als »intelligent« bezeichneten Maschinen begonnen hat.

Maßgeblicher Motor der aktuellen Innovationsdynamik ist eine rasante Zunahme der Rechenkapazität von Computern, gekoppelt an die zunehmende Miniaturisierung, Vernetzung und Verbilligung der relevanten technischen Komponenten. Der Fortschritt im Bereich der Computer- und Informationstechnologie macht es in Zusammenhang mit Erkenntnissen aus dem Bereich der Neurowissenschaften und der KI inzwischen möglich, Maschinen zu entwickeln, welche das Potenzial haben, das Mensch-Technik-Verhältnis ganz neu zu definieren. Zwei unterschiedliche Technologiefelder sind diesbezüglich von besonderer Relevanz, nämlich die Neurotechnologien auf der einen und die autonome Robotik auf der anderen Seite:

- > *Neurotechnologien* umfassen die invasive oder nichtinvasive Kopplung maschineller Systeme an das Gehirn bzw. periphere Nervensystem. Dadurch wird es möglich, neuronale Signalprozesse direkt zu beeinflussen oder aber neuronale Signale über eine technische Schnittstelle auszulesen, mittels maschinellen Lernens zu interpretieren und in maschinelle Aktivität umzusetzen (Steuerung von Neuroprothesen, Computern etc.). Damit eröffnen sich ganz neue Möglichkeiten, körperliche Defizite wie Gehörlosigkeit oder Blindheit technisch zu kompensieren (mithilfe sogenannter Neuroimplantate oder -prothesen) oder Gliedmaßenprothesen anzusteuern. Die Technisierung des Menschen erreicht dadurch eine ganz neue Qualität, da diese Geräte immer häufiger direkt am Gehirn ansetzen, also an der Quelle des menschlichen Bewusstseins und Selbstbewusstseins.



- › *Roboter* sind dank ihres Autonomisierungsgrades bereits seit längerem in der Lage, bislang Menschen vorbehaltene Tätigkeiten auszuüben. Neben dem klassischen und weitverbreiteten Industrieroboter, dessen Verhalten noch in vielerlei Hinsicht determiniert ist, wird in den Laboren von Forschung und Industrie aber auch an neuen Robotertypen geforscht, die zunehmend Tätigkeiten übernehmen sollen, die bisher nur von Menschen durchgeführt werden konnten. Hierzu gehören komplexe Aufgaben, die bezüglich kognitiver Fähigkeiten (Autonomie, Intelligenz, Lernfähigkeit) einen hohen Entwicklungsgrad erfordern. Dadurch eröffnen sich neue Anwendungsfelder, primär in den Bereichen Industrie, Haushalt, Medizin bzw. Pflege sowie beim Militär – konkrete Anwendungen sind etwa Serviceroboter, die (in sehr einfacher Form) bereits im Haushalt oder in der Pflege verbreitet sind. Von traditionellen Maschinen unterscheiden sich solche Maschinen dadurch, dass sie nicht nur die Handlungsfähigkeit des Menschen erweitern, sondern selbst in einem funktionalen Sinne quasi handlungsfähig sind, ja sein müssen, um die ihnen zugewiesenen Aufgaben erledigen zu können.

Im »Zeitalter der Neurotechnologien« bezieht der *Homo faber* somit erstmals seinen inneren Wesenskern in den Technisierungsprozess mit ein (Müller 2010, S. 9), während er sich mit Robotern quasi sein technisches Ebenbild zu schaffen sucht. Verbindendes Kernmerkmal dieser (auf technologischer Ebene) noch weitgehend unverbundenen Entwicklungsstränge ist folglich, dass sie die tradierte anthropologische Abgrenzung zwischen dem schaffenden Menschen einerseits und seinen technischen Hilfsmitteln andererseits grundlegend und jeweils aus unterschiedlichen Richtungen herausfordern. Aus übergeordneter Perspektive zeigt sich eine wechselseitige Dynamik, die im Folgenden als Mensch-Maschine-Entgrenzung bezeichnet wird und futuristisch erscheinende Perspektiven eröffnet.

So bietet der neurotechnologische Zugriff auf das Gehirn nicht nur die Chance, durch Krankheit oder Unfall hervorgerufene Defizite auszugleichen, sondern im Prinzip ist es bereits jetzt möglich, die psychische und körperliche Leistungsfähigkeit des Menschen mittels technischer Erweiterungen, die im oder am Körper getragen werden – sei es in Form von Prothesen, Neuroimplantaten oder Gehirn-Computer-Schnittstellen –, gezielt zu verbessern. Inwiefern eine derartige Verschmelzung des Menschen mit und damit seine steigende Abhängigkeit von der Technik gesellschaftlich wünschenswert und unter ethisch-moralischen Gesichtspunkten vertretbar ist, wird unter dem Schlagwort *Human Enhancement* (HE) bereits seit längerem kontrovers diskutiert. Dass sich die vor allem philosophisch geprägten Fragestellungen vor dem Hintergrund gegenwärtiger Entwicklungen als aktueller denn je erweisen, zeigt u. a. das Beispiel des sogenannten »Cyborgismus«, bei dem es sich um eine Bewegung von Men-



schen handelt, die vor allem mit Technologien zur Sinneserweiterung experimentiert und dabei auch Implantate nutzt. Maschinenmenschen, sogenannte Cyborgs, sind also anscheinend keine rein futuristische Vision mehr. Dies wirft grundlegende Fragen auf: Wo verläuft die Grenze zwischen Mensch und Maschine? Welche Folgen hat es für unsere Gesellschaft, wenn sich solche technischen Erweiterungen, die zunächst nur für einen kleinen Personenkreis bezahlbar sein dürften, durchzusetzen beginnen?

Parallel zu der zunehmenden Technisierung des Menschen ist zu beobachten, dass körperextern agierende Maschinen (zumindest in funktionaler Hinsicht) immer menschenähnlicher werden und in machtvollere Rollen zu schlüpfen beginnen. Schon heute ist maschinelles Lernen anhand großer Datenmengen ein etabliertes Phänomen, das Eingang in verschiedene Alltagsanwendungen gefunden hat und etwa bei Überwachungstechnologien oder bei Suchmaschinen zum Einsatz kommt. Zwar ist der superintelligente Roboter mit einem Ichbewusstsein noch ein unbestimmtes Zukunftsszenario, aber es ist bereits heute ein Punkt im Automatisierungsgrad erreicht, der nicht nur das menschliche Selbstverständnis, sondern auch die Gesellschaft grundlegend revolutionieren könnte. Inwieweit wird z. B. die Autonomie von Menschen infrage gestellt, wenn sich derartige Anwendungen in sensiblen Alltagsbereichen wie der Pflege durchzusetzen beginnen? Neben ethischen wirft diese Entwicklung insbesondere diffizile rechtliche Fragen auf, vor allem im Hinblick auf Sicherheit, Haftung und Datenschutz. Wie können die technischen Systeme und die in ihnen gespeicherten Daten juristisch vor Angriffen durch Dritte und Missbrauch geschützt werden? Wer ist im Falle von Fehlfunktionen für die entstehenden Schäden verantwortlich?

Zielsetzung und Zuschnitt des Berichts

Eine besondere Relevanz gewinnt die Entgrenzungsthematik vor dem Hintergrund jüngst gestarteter Großprojekte in der EU (»Human Brain Project«) und den USA (»Brain Initiative«), in denen die Funktionsprinzipien des Gehirns mithilfe von Computern modelliert und simuliert werden sollen, nicht zuletzt mit dem Ziel, neue Anwendungsfelder im Schnittfeld von Robotik und Neurowissenschaften bzw. -technologien zu erschließen. Angesichts der vielschichtigen Tragweite dieser Entwicklungen erscheint eine rechtzeitige Auseinandersetzung mit den zugrundeliegenden wissenschaftlich-technischen Fragen sowie den daraus erwachsenden Anwendungspotenzialen und Nebenfolgen also überaus sinnvoll und notwendig – umso mehr, als sich der technologische Fortschritt weitgehend schleichend vollzieht und unumkehrbare gesellschaftliche Veränderungen einleiten könnte. Vor diesem Hintergrund wurde das TAB vom zuständigen Ausschuss für Bildung, Forschung und Technikfolgenabschätzung



damit beauftragt, eine Untersuchung zum Thema »Mensch-Maschine-Entgrenzungen« durchzuführen.

Eine Abschätzung möglicher Technikfolgen in diesem Bereich steht jedoch vor mehreren Schwierigkeiten, wie sie für TA-Projekte typisch sind, die sich mit neuen und emergierenden Wissenschafts- und Technikfeldern befassen (sogenannte »New and Emerging Science and Technology« [NEST]), zu denen auch die hier angesprochenen gehören. Verantwortlich dafür sind im Wesentlichen die drei folgenden Merkmale der NEST:

- › Sie befinden sich noch in einem sehr frühen, grundlegenden Entwicklungsstadium, sodass Prognosen bezüglich realisierbarer Anwendungspotenziale und deren gesellschaftlicher Implikationen nur bedingt möglich sind – zu groß erscheint hierfür der »Möglichkeitsraum des Zukünftigen« (Grunwald 2015, S.66).
- › Prognosen bezüglich der weiteren Entwicklungsdynamik werden auch dadurch behindert, dass es sich in der Regel nicht nur um äußerst dynamische, sondern auch um disziplinär weitverzweigte Forschungs- und Entwicklungsfelder handelt. Hinzu kommt, dass sich maßgebliche Innovationsimpulse nicht selten durch unerwartete Konvergenzprozesse an der Schnittstelle unterschiedlicher Wissenschafts- und Technologiebereiche ergeben. Zu nennen sind vor allem die Nano-, Bio-, und Informationstechnologien sowie die Neuro- und Kognitionswissenschaften (TAB 2008).
- › Die unscharfen Konturen der Wissenschafts- und Technikfelder werden durch öffentliche Debatten konterkariert, die um futuristische, aber dennoch teils erstaunlich präzise formulierte Technikvisionen kreisen. Was die gesellschaftliche Wünschbarkeit dieser weitreichenden Zukunftserwartungen angeht, bestehen häufig sehr kontroverse Einschätzungen. Insgesamt resultiert so eine diskursive Gemengelage, die eine nüchterne Betrachtung von Chancen und Risiken erschwert.

Um zu einer seriösen Einschätzung derselben zu kommen, ist es deshalb wichtig, zunächst den aktuellen Stand von Forschung, Entwicklung und Anwendung zu beschreiben, um auf dieser Basis dann die tatsächlich relevanten (anstelle rein hypothetischer und spekulativer) Herausforderungen zu eruieren. Dies ist grob umrissen die wesentliche Zielsetzung des vorliegenden Sachstandsberichts. Damit soll eine solide Basis geschaffen werden, um die komplexe Entgrenzungsdynamik, ihre wissenschaftlich-technischen Perspektiven und gesellschaftlichen Implikationen im weiteren Projektverlauf anhand konkreter Anwendungsfelder vertiefen und konkretisieren zu können.

Der Sachstandsbericht hat folgende Struktur: In Kapitel II wird der aktuelle visionäre Diskurs zu Mensch-Maschine-Entgrenzungen durch Neurotechnologie, KI und Robotik skizziert, dabei werden auch historisch-kulturelle und ge-

sellschaftliche Kontexte der Zukunftsvisionen beleuchtet. Vor diesem Hintergrund werden dann in Kapitel III der aktuelle Stand von Forschung und Entwicklung in den Bereichen Neurotechnologien sowie der autonomen Robotik dargestellt, bevor in den Schlussfolgerungen (Kap. IV) die Realisierbarkeit der Visionen vorläufig bewertet und relevante normative Fragen umrissen werden.

In Auftrag gegebene Gutachten

Der vorliegende Bericht basiert auf folgenden Kurzgutachten, die im Rahmen des Projekts vergeben wurden:

- > Das Gehirn im Mittelpunkt. Neuroelektrische Schnittstellen schließen den Regelkreis. Prof. Dr. Ulrich Hofmann, Freiburg
- > Autonome Robotik, Maschinelles Lernen & Künstliche Intelligenz. Fraunhofer-Institut für Intelligente Analyse- und Informationssysteme (Autor: Prof. Dr. Sven Behnke), Sankt Augustin
- > Entwicklungen in den Bereichen autonome Robotik, maschinelles Lernen & künstliche Intelligenz. Prof. Dr. Stefan Schaal, Santa Monica, USA
- > Vision Assessment: Mensch-Maschine-Entgrenzungen. Zwischen künstlicher Intelligenz und Human Enhancement. Fraunhofer-Institut für System- und Innovationsforschung (Autoren: Dr. Bruno Gransche, Dr. Philine Warnke), Karlsruhe

Der Gutachterin und den Gutachtern sei für die fruchtbare und engagierte Zusammenarbeit herzlich gedankt. Ein besonderer Dank geht auch an Hannah Weinhardt für eine Auswertung und Zusammenfassung der TA-Debatten zum Thema, an Matthias Körber für ergänzende Recherchen und an Dr. Arnold Sauter, Knud Böhle und Reinhard Heil für die kritische Durchsicht des Manuskripts sowie an Marion Birner und Brigitta-Ulrike Goelsdorf für die Aufbereitung der Abbildungen und die Erstellung des Endlayouts.



Zukunftsvisionen

II.

Die aktuellen Diskussionen über die in diesem Bericht behandelten Technologien werden in hohem Maße durch weitreichende Zukunftsvisionen beeinflusst. Die Visionen wiederum fügen sich ein in bereits seit langer Zeit existierende, z. T. überlappende Diskurse zur Technisierung des menschlichen Körpers einerseits und zu autonomen oder intelligenten Maschinen andererseits. Diese Diskurse sind kulturell sehr wirkmächtig und haben ganze Genres, wie z. B. die Science-Fiction (in der Literatur, im Film und Spielebereich), maßgeblich geprägt. Zudem – und auch: daher – können die gegenwärtigen Diskussionen über HE und KI als zwei Paradebeispiele für die in den letzten Jahren zunehmend vertretene Ansicht gelten, dass bei der Ab- und Einschätzung neuer oder in der Entstehung befindlicher (emergierender) Technologien zumeist auch eine systematische Untersuchung von Zukunftsvisionen – ggf. einschließlich sehr weitreichender Visionen – angezeigt ist. Im Fall der Debatte über HE, das im Deutschen zumeist entweder als »Steigerung der menschlichen Leistungsfähigkeit« oder als »Verbesserung des Menschen« übersetzt wird, ist dabei eine Beschränkung auf technologische Eingriffe in den Körper und die Einnahme von Substanzen angezeigt (Coenen et al. 2009), da bei breiteren Definitionen fast jedwede Techniknutzung als HE gelten würde, wodurch die Debatte ihren spezifischen Sinn verlöre.

Im gegebenen Rahmen können keine umfassende Visionsanalyse und -bewertung für die in diesem Bericht behandelten naturwissenschaftlich-technischen Entwicklungen und Themen vorgenommen werden. Es sollen aber zum einen – nach einer kurzen Einführung zur Bedeutung von Zukunftsvisionen in der Technikfolgenabschätzung und einer knappen Skizze der Konturen des visionären Diskurses zu Mensch-Maschine-Entgrenzungen (Kap. II.1) – die zentralen Ideen und Zukunftsbilder in den visionären Diskussionen aus historischer Perspektive dargelegt werden (Kap. II.2). Eine solche historische Perspektive ist angezeigt, da die aktuellen Diskussionen ganz wesentlich auf älteren Ideen und Debatten basieren. Zum anderen soll beispielhaft auf die aktuellen visionären Diskussionen sowohl zu HE und Cyborgs als auch zu KI eingegangen werden (Kap. II.3). Hierbei geht es insbesondere auch darum, die Akteursnetzwerke darzustellen, die diese Diskussionen maßgeblich vorantreiben, da ohne eine solche Darstellung die Dynamik vor allem der massenmedialen Debatten nicht zu verstehen ist. In einem Fazit (Kap. II.4) werden dann einige Schlussfolgerungen aus den vorherigen Teilkapiteln gezogen und die Ergebnisse in einen breiteren Kontext eingeordnet und bewertet.

Zur Analyse des visionären Diskurses

1.

Visionäres Denken wird heute in Politik, Wirtschaft, Wissenschaft und anderen Bereichen oft eingefordert. Unternehmen und andere Organisationen formulieren, z. B. in sogenannten »Vision Statements«, ihre »Vision« für die Zukunft. Ungeachtet des bekannten Ausspruchs des unlängst verstorbenen ehemaligen Bundeskanzlers Helmut Schmidt – wer eine Vision habe, solle zum Arzt gehen –, wird auch von der Politik erwartet, dass sie Visionen in diesem Sinne entwickelt und die Folgen potenziell einschneidender Entwicklungen längerfristig zu antizipieren versucht.

Beides gilt natürlich besonders für die Forschungs- und Technologiepolitik, in der durch die Technikfolgenabschätzung (TA) – und durch engverwandte oder weitgehend identische Forschungsbereiche wie Foresight und die Innovations- und Technikanalyse – politischen Institutionen Orientierungs- und Handlungswissen für antizipatorische Governance bereitgestellt wird. Zunehmend wird dabei auch Wissen zu den jeweils relevanten Zukunftsvisionen einbezogen. In den letzten Jahrzehnten und insbesondere im 21. Jahrhundert hat das Interesse an den Funktionen und Bedeutungen von Zukunftsvisionen im Diskurs zu forschungs- und technologiepolitischen Fragen deutlich zugenommen. Auch in der wissenschaftlichen Politikberatung, nicht zuletzt für den Deutschen Bundestag (z. B. Gransche/Warnke 2014; Paschen et al. 2004; TAB 2008), werden nun häufiger Zukunftsvisionen sowie deren Erzeugung und Nutzung analysiert (z. B. Berloznik et al. 2006; Coenen et al. 2009; van Est/Stemerding 2012; Felt/Wynne 2007). Dabei kann die interdisziplinäre Forschung, die in Technikfolgenabschätzung und verwandten Feldern geleistet wird (z. B. Böhle/Bopp 2014; Ferrari et al. 2012), auch auf Erkenntnisse klassischer Disziplinen wie der Geschichtswissenschaft (McCray 2012), der Philosophie (z. B. Kaiser 2015; Nordmann 2007; Nordmann 2013) und der Soziologie (z. B. Jasanoff/Kim 2009; Selin 2008) zurückgreifen (siehe dazu Coenen/Simakova 2013; Grunwald 2013; Schneider/Lösch 2015).

Wie lässt sich die Rolle von Zukunftsvisionen im Diskurs über Naturwissenschaft und Technik charakterisieren? In einer von acatech – Deutsche Akademie der Technikwissenschaften herausgegebenen Studie heißt es, dass »Zukunftsvorstellungen«, die in Wissenschaft, Kunst und Kultur oder Journalismus artikuliert werden, eine »entscheidende Rolle in gesellschaftlichen Technikdebatten« spielten (acatech 2012, S.6). Zukunftsvorstellungen erfüllten »in der Forschungspolitik die Funktion der Legitimation öffentlicher Förderprogramme« und hätten daher »maßgeblichen Anteil an der Gestaltung der Agenda in Forschung und technischer Entwicklung«. Vor allem aber seien solche Vorstellungen relevant für gesamtgesellschaftliche »Diskussionen über die Frage, mit welcher Technik wir als Gesellschaft zukünftig leben wollen«. Die unterschiedli-



chen Zukunftsvorstellungen prägten aber nicht nur die gesellschaftlichen und politischen Debatten, sondern auch »das konkrete Entwicklungshandeln der Ingenieure« und somit auch auf dieser Ebene »die Ausgestaltung zukünftiger technischer Systeme und deren Nutzungsbedingungen« (acatech 2012, S.6). Im Folgenden werden nun in diesem Sinne Zukunftsvisionen zu Technologien der Mensch-Maschine-Entgrenzung vorgestellt und diskutiert.

Betrachtet man allein die naturwissenschaftlich-technischen Grundlagen und Anknüpfungspunkte der visionären Diskussionen, liegt es nicht unbedingt nahe, neurotechnologische HE-Visionen und Zukunftsvisionen zu KI und autonomen Robotern gemeinsam zu betrachten. Zwar existieren – insbesondere im Feld avancierter Prothetik – interessante und vielversprechende technologische Konvergenzen, dennoch sind Neurotechnologien einerseits und KI (einschließlich der auf KI setzenden Robotik) andererseits klar voneinander abgrenzbare Felder der Forschung und Technologieentwicklung.

Dasselbe gilt für die populärsten Visionen zu diesen Feldern, die sich zu meist klar voneinander unterscheiden lassen. Beispielhaft können hier Cyborgs und Androiden genannt werden: Handelt es sich bei den Erstgenannten um Wesen, in denen der menschliche Körper mit Technik verschmolzen ist, handelt es sich bei den Letztgenannten um intelligente und Menschen täuschend ähnlich sehende Roboter. In Science-Fiction wie Massenmedien kommen zu diesem Duo oft noch hochintelligente KI-Entitäten hinzu, die nicht in humanoiden Körpern residieren, aber über Selbstbewusstsein verfügen und mit Menschen – oft aus einer überlegenen Position heraus – interagieren. Eine in letzter Zeit – u. a. durch Wally Pfisters Film »Transcendence« (2014) – bekannter werdende Mischform ist die Vision eines Hochladens des »Geistes« eines menschlichen Individuums auf Maschinen (das sogenannte Mind Uploading), woraus dann – so die Annahme – ein in den Maschinen oder sogar in digitalen Netzwerken existierendes, sich seiner selbst bewusstes Wesen menschlicher »Herkunft« entstehen könnte (siehe dazu Seung 2013), das wiederum dazu in der Lage wäre, mit anderen Individuen gleichsam technotelepathisch zu kommunizieren. Visionen dieser Art werden seit den 2000er Jahren besonders öffentlichkeitswirksam von dem erfolgreichen Ingenieur, Unternehmer und prominentesten Transhumanisten Ray Kurzweil (2005) propagiert.

In den Massenmedien und auch in der Science-Fiction tauchen diese visionären Denkfiguren häufig zusammen auf und werden miteinander vermengt. Daneben lassen sich weitere Gründe dafür nennen, warum beide visionären Diskussionen in der Zusammenschau betrachtet werden sollten:

- › Die naturwissenschaftlich-technischen Entwicklungen, die oft als Ausgangspunkte der Diskussionen dienen, lassen sich aus anthropologischer und technikphilosophischer Sicht als Vorboten oder erste Anzeichen einer umfassenden Mensch-Maschine-Entgrenzung begreifen, in der eine Tech-



nisierung des menschlichen Körpers – vorbereitet durch eine Technisierung des Menschenbildes (Grunwald 2009) – und die Erschaffung anthropomorpher oder den Menschen bzw. menschliche Fähigkeiten imitierender Technik Hand in Hand gehen (für eine auch der Politikberatung dienende Technikfolgenabschätzung dazu van Est/Stemerding 2012)

- › Die Zusammenschau der beiden Diskussionen trägt auch dem Umstand Rechnung, dass bereits seit dem späten 20. Jahrhundert, deutlich verstärkt aber seit den 2000er Jahren eine umfassende Konvergenz – wenn nicht gar gezielte Zusammenführung – sehr weitreichender Technikvisionen aus verschiedenen Feldern (z. B. Life Sciences und Biotechnologie; Informatik und Computertechnik; Hirnforschung und Neurotechnologie; Nanotechnologie; Militärforschung; Weltraumforschung) stattfindet oder zumindest weithin postuliert wird (TAB 2008). Diese Diskussionen stehen im Zeichen einer Renaissance der alten, oft als übertrieben oder gar (als für das Forschungsfeld) verhängnisvoll charakterisierten KI-Euphorie, die sich weniger im vorsichtig gewordenen Mainstream der KI-Forschung als an deren Rändern und vor allem in anderen Bereichen äußert.
- › Den weltanschaulichen Rahmen für die Konvergenz (oder auch Zusammenführung) der Zukunftsvisionen zu verschiedenen Technikfeldern bietet immer öfter der sogenannte Transhumanismus, dem es erklärtermaßen darum geht, die *Conditio humana* radikal zu verändern und z. B. das Altern, die Grenzen menschlicher und künstlicher Intelligenz, eine nicht selbstgewählte Psyche, menschliches Leiden und unsere Beschränkung auf den Planeten Erde zu überwinden (Heil 2010a; Schneider 2009, S. 97). Während die Ideengeschichte des Transhumanismus mindestens bis in die 1920er Jahre zurückreicht, entstand eine organisierte transhumanistische Bewegung erst in den 1980er Jahren in den USA. Sie hat sich aber seit Ende der 1990er Jahre in verschiedenen Ländern akademisch, zivilgesellschaftlich und politisch etabliert, was sich in letzter Zeit sogar in diversen Parteigründungen niedergeschlagen hat (Wagner 2015).
- › Auf der Akteursebene sind hinsichtlich der Visionsproduzenten und -verbreiter und ihrer Unterstützer erhebliche Überlappungen der beiden Felder festzustellen (Kap. II.3).
- › Bemerkenswert ist u. a., dass zwar – insbesondere dezidiert transhumanistische – Teile dieser Netzwerke bereits seit den 1970er und 1980er Jahren bestehen (Schummer 2009) und schon z. B. den visionären Diskurs zur Nanotechnologie stark geprägt haben (Paschen et al. 2004; Simakova/Coenen 2013), aber im neuen Jahrhundert sich nicht nur der Organisationsgrad und der akademische Einfluss der transhumanistischen Bewegung erheblich erhöht haben, sondern sich zudem auch – neben einflussreichen Industriellen – (weitere) prominente und renommierte Naturwis-

senschaftler und Ingenieure sowie u. a. auch einige namhafte Sozial- und Geisteswissenschaftler, Künstler und Journalisten an den Diskussionen zu beteiligen begonnen haben.

Kulturelle Wurzeln und Ideentraditionen

2.

Im gegebenen Rahmen kann nur eine sehr grobe Skizze kultureller Wurzeln und Ideentraditionen erfolgen, die in diesem Zusammenhang bedeutsam erscheinen. Eine solche Skizze ist aber notwendig, da sich der Sinn der aktuellen Diskussionen ohne Berücksichtigung ihres historisch-kulturellen Hintergrunds kaum erschließen lässt.¹

Künstliche Menschen, Automaten und Roboter

2.1

Ein naheliegender Ausgangspunkt für diese ideen- und kulturgeschichtliche Skizze ist die historisch sehr weit zurückreichende Faszination durch die Vorstellung künstlicher Menschen (Tabbert 2004; siehe auch Meyer-Drawe 1996): Hinsichtlich dieser ist – neben den für das Thema dieses Berichts weniger relevanten Visionen der Alchemie und Religion (Golem, Homunculus) – auf die bereits für die Antike nachweisbare Entwicklungstätigkeit im Bereich der Automaten hinzuweisen, die dann in der Renaissance zunehmend auf humanoide Artefakte ausgeweitet wurde. Das Beispiel des berühmtesten Automatenkonstruktors seiner Zeit, Jacques de Vaucanson (1709–1782) – der nicht nur einen für die Entwicklung der modernen Textilindustrie grundlegenden automatischen Webstuhl entwickelte, sondern besonderes Interesse an der Erschaffung möglichst lebensähnlicher Automaten in Menschen- und Tierform (z. B. Flötenspieler bzw. Ente) hatte –, zeigt, dass spielerische und nützliche Aspekte der Automatenkonstruktion in dieser Zeit durchaus miteinander verschränkt waren.

Die Automaten Vaucansons und seiner Zeitgenossen sind zudem nicht nur von Interesse, weil sie oft Produkte realer Technikentwicklung waren (und keineswegs nur Pseudoautomaten wie der immens populäre »Schachtürke«, bei dem die Bedienung geheim von einem kleinwüchsigen Menschen übernommen wurde). Sie zeigen zudem, dass das visionäre Ziel, künstliche Lebewesen (Menschen oder Tiere) herzustellen, auch weitverbreitete philosophische oder wissenschaftliche Grundannahmen der jeweiligen Zeit widerspiegeln kann. Die

¹ Aus Platzgründen ausgespart oder nur am Rande behandelt werden allerdings einige durchaus relevante Themen, wie z. B. das erhebliche Ausmaß der Prägung relevanter Diskurse durch die Militär- und die Weltraumforschung (z. B. Coenen et al. 2009; TAB 2008). Überdies fokussieren die Ausführungen auf westliche Ideentraditionen.



z. B. für die moderne Medizin und das neuzeitliche europäische Verständnis des Menschen zentralen mechanistischen Auffassungen, wie sie von René Descartes (1596–1650), Julien Offray de La Mettrie (1709–1751) und anderen entwickelt wurden, fanden einen Ausdruck in den Automaten des 18. Jahrhunderts. Dementsprechend äußerte sich die insbesondere ab Anfang des 19. Jahrhunderts zunehmende Fortschrittsskepsis ebenfalls in Visionen menschenähnlicher künstlicher Wesen. Zu den Hintergründen dieses wachsenden Zweifels am Fortschrittsdenken zählten das Sichtbarwerden negativer Folgen der industriellen Revolution und die Enttäuschung aufklärerischer Hoffnungen im Zuge der Französischen Revolution. Bei den Visionen handelte es sich dementsprechend oft um Schreckensvisionen, wie z. B. in Mary Shelleys Roman »Frankenstein or The Modern Prometheus« (1818).

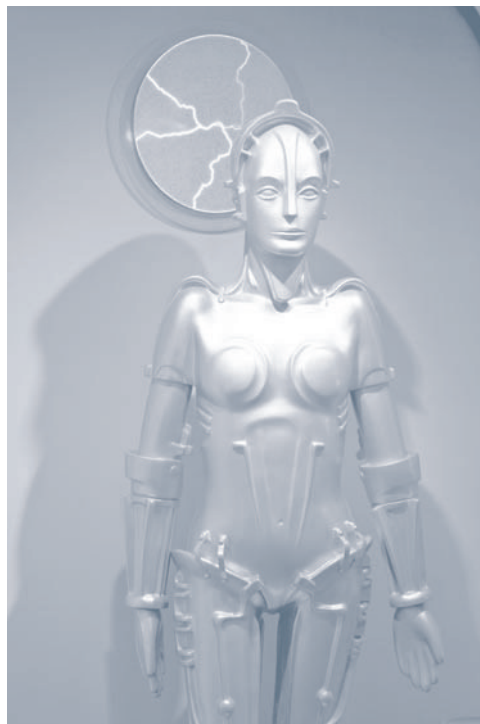
Obwohl ab der zweiten Hälfte des 19. Jahrhunderts und dann verstärkt im 20. Jahrhundert biologische Visionen künstlicher oder technisch massiv umgeformter Menschen, einschließlich der neuen Cyborgvisionen, an Relevanz gewannen, blieben die Automatenvisionen von einiger Bedeutung. Zwar handelt Karel Čapeks literarisches Werk »R.U.R.« (Rossum's Universal Robots) aus dem Jahr 1920 – in dem der (von einem slawischen Wort für harte Arbeit abgeleitete) Begriff »Roboter« geprägt wird – von synthetisch erzeugten biologischen Wesen und nicht von den bald weltweit populären, als »Roboter« bezeichneten und damals noch visionären Automaten (z. B. durch Fritz Langs epochalen Film »Metropolis« von 1927, Abb. II.1). Dennoch ist Čapeks dystopisches Drama ein bedeutsamer Ausdruck der mindestens seit dem 19. Jahrhundert bekannten Vorstellung einer wissenschaftlich-technisch erzeugten Sklavenpopulation, die der Menschheit zu Diensten sein wird. Noch heute versinnbildlichen weitreichende Robotisierungsvisionen eine Antriebskraft der gesamten westlichen, wenn nicht humanen Technikfortschrittsgeschichte: die Hoffnung auf eine Befreiung des Menschen von dem (im biblischen Sinn) nachparadiesischen Zwang zur Arbeit.

Durch die in den 1930er bzw. 1940er Jahren erfolgende Entwicklung des Computers und der Kybernetik, eines einflussreichen interdisziplinären Forschungsansatzes, wurden die heutigen Visionen zu KI und Robotik, aber auch zu HE, maßgeblich geprägt. Aus Kybernetik, Computerwissenschaft und Informatik entwickelte sich die KI-Forschung, die insbesondere anfangs stark durch die sogenannte »starke« KI-These geprägt wurde (Kap III.2). Durchaus im Sinne der alten aufklärerischen Automatentradiation und der ihr zugrundeliegenden philosophischen und wissenschaftlichen Annahmen – zu denen in gewisser Hinsicht das Begreifen auch des Menschen als Maschine zählte –, geht es der »starken« KI darum, eine menschenähnliche, den Menschen aber dann intellektuell überflügelnde Intelligenz zu schaffen. In den bedeutenden futurologischen und Science-Fiction-Werken des Autors Stanislaw Lem wurden u. a. auch mögliche Implikationen einer zukünftigen »starken« KI vielschichtig the-

matisiert. Die wohl bekannteste Schreckensvision findet sich in Stanley Kubricks Film »2001: A Space Odyssey« (1968), der auf einem Werk des bedeutenden Technikvisionärs Arthur C. Clarke basiert sowie in Zusammenarbeit mit diesem realisiert wurde und in dem der hochintelligente Bordcomputer »HAL 9000« versucht, die menschliche Besatzung eines Raumschiffs zu töten.

Abb. II.1

Der Maschinenmensch »Maria« aus Metropolis (Replikat)



Quelle: Jiuguang Wang/Wikimedia Commons/CC-BY-SA-2.0

Eine Schlüsselfigur in der Ideen-, Forschungs- und Entwicklungstradition der »starken« KI ist Marvin Minsky, ein Mitbegründer der KI-Forschung. An ihm wird häufig auch eine zentrale Kritik an der »starken« KI festgemacht, dass deren Befürworter nämlich verschiedentlich viel zu weitreichende Erwartungen in die Welt gesetzt und damit die beiden sogenannten »KI-Winter« in der zweiten Hälfte der 1970er Jahre und in den Jahren um 1990 (Perioden weitverbreiteter Enttäuschung über das Forschungsfeld und entsprechender Zurückhaltung bei dessen Förderung) maßgeblich mitverursacht hätten (Kap. III.2.1). Tatsächlich gehört Minsky – wie z. B. auch der Robotiker Hans Moravec oder der Nanofuturist Eric Drexler (Coenen et al. 2009; Drexler 1986; Paschen et al. 2004) – zu denjenigen Forschern bzw. »Visioneers« (Visionsingenieuren; McCray 2012), die mit gewagten, z. T. auch in populärwissenschaftlichen Werken verbreiteten Zukunftsvisionen das öffentliche Bild des Forschungsfeldes sowie anderer Felder massiv beein-



flusst haben. Im Fall von Minsky, Moravec, Drexler und anderen sind diese Visionen zudem stark transhumanistisch geprägt und enthalten auch sehr weitreichende Erwartungen wie die einer Abschaffung des Todes oder Ablösung bzw. Ersetzung der Menschheit durch KI (Kap. II.3.1).

Eine zentrale Bedeutung hat in diesen Diskussionen die sogenannte Singularitätshypothese. Sie geht zurück auf ein 1965 vorgestelltes Konzept der »Intelligenzexplosion« des Mathematikers Irving J. Good (1965), das besagt, dass eine dem Menschen überlegene KI eine weitere, noch intelligentere KI erzeugen können wird, die dann wiederum eine noch intelligentere KI entwickelt usw., wodurch es zu der besagten Intelligenzexplosion kommen werde.² Das Überleben der Menschheit hänge von der baldigen Konstruktion einer solchen Maschine ab. Die Erfindung einer »ultraintelligenten« Maschine werde zudem womöglich die letzte menschliche Erfindung sein, da solche Maschinen, vorausgesetzt, sie ließen sich unter Kontrolle halten, dann zum Wohle der Menschheit auch die Erfindungstätigkeit übernehmen könnten.

Durch den Mathematiker, Computerwissenschaftler und Science-Fiction-Autor Vernor Vinge wurde dann im Jahr 1993 ein Konzept der »Singularität« vorgestellt,³ in dem die bevorstehende Heraufkunft einer dem Menschen überlegenen KI als das bedeutendste Ereignis der menschlichen Geschichte fungiert, und das auf folgenden Annahmen basiert: Es wird ichbewusste und übermenschlich intelligente Computer geben, große Computernetzwerke samt ihrer menschlichen Nutzer werden sich dann zu einer übermenschlich intelligenten Entität entwickeln und die biologischen Wissenschaften könnten zudem Mittel zur Verfügung stellen, den natürlichen menschlichen Intellekt zu steigern (biologisches Enhancement). Ausgebaut und popularisiert wurde die Singularitätshypothese dann vor allem durch Ray Kurzweil (2005).

Das die aktuelle öffentliche Diskussion über KI prägende Motiv einer zukünftigen Konkurrenz zwischen KI und der Menschheit – bei der Letztere unter Umständen verklavt werden könnte – ist seit Mitte der 1980er Jahre auch in der populären Kultur weithin bekannt. Bereits im Jahr 1984 hatte James Camerons Film »The Terminator« eine Zukunft ausgemalt, in der intelligente Kampfrobooter Krieg gegen ihre menschlichen Erschaffer führen. Die anhaltende Faszination durch dieses Thema zeigen der ähnliche Erfolg des Films »The Matrix« (1999) der Wachowski-Geschwister – dessen Handlung die Idee einer (nur wenigen Menschen bewussten) Verklavung der Menschheit durch KI-Entitäten zugrunde liegt – sowie der Umstand, dass beide Filme zu erfolgreichen Trilogien ausgebaut wurden. Vor diesem Hintergrund ist verständlich, wieso die sogenannten Asimovschen Robotergesetze des Wissenschaftlers und Schriftstellers Isaac Asimov

2 Good war auch wissenschaftlicher Berater Stanley Kubricks für dessen Film »2001: Space Odyssey«.

3 <http://mindstalk.net/vinge/vinge-sing.html> (15.12.2015)

(1919–1992) sowohl in ethischen als auch in technischen Diskussionen zur Robotik bis heute eine einflussreiche Rolle spielen – weit über die Science-Fiction hinaus. In seinen literarischen Werken entwarf Asimov Szenarien eines zukünftigen lebensweltlichen Miteinanders von Menschen und Robotern, in denen der Schutz des Menschen ganz im Vordergrund stand. Roboter sind demzufolge intelligente Automaten, für die »Gesetze« zu gelten haben (z. B. »Ein Roboter darf keinen Menschen verletzen oder durch Untätigkeit zu Schaden kommen lassen.«).

Transhumanismus, Cyborgs und Human Enhancement 2.2

Auch die aktuellen Diskussionen über ein HE mit neurotechnologischen Mitteln haben eine längere Vorgeschichte. Für diese waren u. a. verschiedene naturwissenschaftlich-technische Fortschritte prägend, vor allem die seit dem 17. Jahrhundert gewonnenen Erkenntnisse über die Rolle der Elektrizität im menschlichen Nervensystem – die dann auch verschiedentlich zu Hoffnungen führten, Leben mithilfe elektrischer Stimulation zu schaffen – sowie der Aufstieg des biologischen Evolutionsdenkens im 19. Jahrhundert insbesondere im Anschluss an Charles Darwin.

Ein grundlegendes Element dieser Vorgeschichte ist die Bestimmung des Menschen als ein Mängelwesen, die im deutschsprachigen Raum z. B. von Johann Gottfried Herder, Ernst Kapp, Sigmund Freud und der modernen philosophischen Anthropologie (Max Scheler, Helmuth Plessner, Arnold Gehlen etc.) vorgenommen wurde. In diesem anthropologischen Ansatz kann Technik als Organprojektion oder -ersatz, als eine Körpererweiterung erscheinen – oder auch der Mensch insgesamt, in den Worten Freuds, als »Prothesengott«. Nicht zuletzt durch den infolge des modernen Kriegswesens und insbesondere des Ersten Weltkriegs stark ansteigenden Bedarfs an Prothesen und einer Eingliederung von Verehrten ins industrielle Arbeitsleben gewinnen Visionen und Experimente zu Mensch-Technik-Kopplungen im 20. Jahrhundert erheblich an Bedeutung.

Nahezu alle aktuellen Diskussionen über HE, einschließlich derer über neurotechnologisches Enhancement, wurzeln zudem in den transhumanistischen Zukunftsvisionen, die infolge des Aufstiegs des biologischen evolutionären Denkens im 19. Jahrhundert und speziell des Darwinismus entstanden sind (z. B. Coenen 2013; Heil 2010b; Saage 2011; Tirosh-Samuels 2012). Schlüsselfiguren waren hier bekannte Naturwissenschaftler oder Schriftsteller wie Herbert G. Wells, John D. Bernal (Abb. II.2), John B. S. Haldane und Julian Huxley, dessen Bruder Aldous auf die frühen transhumanistischen Visionen der Ersteren mit dem berühmten Roman »Brave New World« (1932) reagierte, auf den wiederum noch heute oft – insbesondere von konservativen HE-Kritikern – in Diskussionen vor allem über gentechnisches HE Bezug genommen wird.



Hinsichtlich der in diesem Bericht im Mittelpunkt stehenden Frage der Mensch-Maschine-Entgrenzung verdient von den genannten Transhumanismuspionieren die meiste Beachtung der brillante Naturwissenschaftler John D. Bernal. In seinem erstmals im Jahr 1929 erschienenen, bemerkenswerten futurologischen Essay »The World, the Flesh and the Devil. An Inquiry into the Future of the Three Enemies of the Rational Soul« (Bernal 1970) – der als ein grundlegendes Werk auch für die Visionen der Science-Fiction gelten kann (Slusser 2009) – finden sich u. a. auch Vorläufer heutiger Mind-Uploading- und Technotelepathievisionen sowie vor allem der Vision einer massiven Cyborgisierung des menschlichen Körpers auf Basis neuroelektrischer Schnittstellen und diese nutzender, avancierter Prothetik.⁴

Ein zentrales Szenario seines Essays sei kurz resümiert, weil es prädigitale Fassungen der Visionen der Cyborgisierung, des Mind Uploading, der telepathischen Kommunikation mittels Technik und der Entstehung einer nichtbiologischen Intelligenz auf exemplarische Weise kombiniert: Letztlich, so Bernal, zähle am Menschen nur das Gehirn. Über ein mit ausreichend frischem Blut versorgtes Gehirn zu verfügen, bedeute »am Leben zu sein – zu denken«. So werde es zwar eine chirurgische Herausforderung, aber möglich sein, Nerven permanent mit einer Maschine zu verbinden, die entweder Botschaften an die Nerven senden oder von diesen erhalten kann. Im Anschluss an eine solche Pioniertat würde ein mechanischer Mensch entstehen, zunächst in Form von Gehirnen in Zylindern, dann als avanciertes Mensch-Maschine-Wesen mit verschiedenen Neuroprothesen, die dann auch mit zusätzlichen sensorischen und motorischen Mechanismen ausgestattet würden sowie andere Maschinen aus Entfernung kontrollieren könnten.

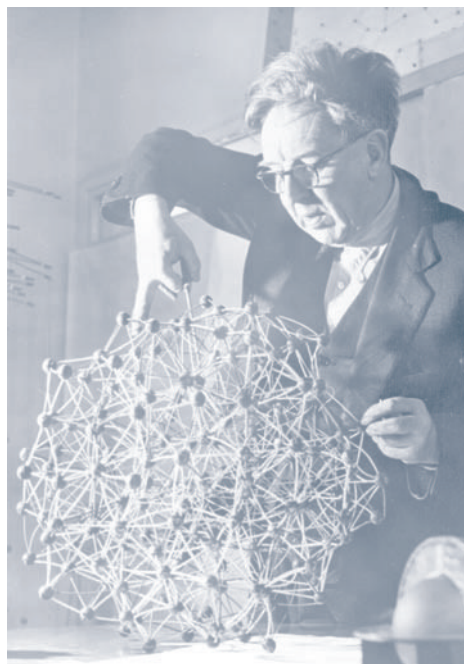
Dieser »neue Mensch« möge auf den ersten Blick wie eine seltsame, monströse und inhumane Kreatur erscheinen, sei aber bloß die logische Folge des

4 Dieses frühe transhumanistische Denken stand wiederum im Kontext eines breiteren Diskurses über als fundamental wahrgenommene Änderungen im Mensch-Maschine-Verhältnis. In diesem Diskurs wurden z. T. die Eingliederung des Menschen in den maschinellen industriellen Arbeitsprozess und sogar die allgemeine lebensweltliche Unterordnung des Menschen unter die Maschine enthusiastisch begrüßt. Besonders einflussreich waren hier der italienische Futurismus (»Manifeste du Futurisme«, 1909, von Filippo Tommaso Marinetti) und, jenseits der Kunst, der Taylorismus, der auch mittels experimenteller Forschung zur Mensch-Technik-Kopplung (z. B. durch Alexei K. Gastew in der Sowjetunion) auf industrielle Prozessoptimierung abzielte. Nicht nur in Deutschland war in diesen Jahrzehnten die Prothetik z. T. darauf ausgerichtet, künstliche Gliedmaßen zu schaffen, die es Verletzten erlauben sollten, durch besondere Funktionalitäten in der Fabrikarbeit beruflich zu reüssieren. So wurden künstliche Glieder entwickelt, die durch auswechselbare Teile sowohl als »Gebrauchsarme für das tägliche Leben« als auch als »Arbeitsarme für den Beruf« dienen konnten (Harrasser 2009). Die wohl bis heute bekannteste Kritik am Taylorismus und verwandten Strömungen ist Charlie Chaplins satirischer Spielfilm »Modern Times« (1936).

heutigen Menschentyps. Eine solche Entwicklung werde, insbesondere auch aus psychologischen Gründen, Zeit brauchen. Aber der »mechanische Mensch«, anscheinend ein Bruch in der organischen Evolution, stehe tatsächlich in der Tradition ständiger Evolution. Deren weiterer Fortgang führt in Bernals Vision dann zunächst von der ekstatischen elektrischen Vernetzung der einzelnen Gehirne zur Herausbildung von multiplen Individuen (bzw. zusammengesetzten Geistern), die sich – aufgrund der Möglichkeit, nach einigen Jahrhunderten absterbende Gehirne unter Beibehaltung der Bewusstseinsinhalte durch neue zu ersetzen – subjektiv als unsterblich und insgesamt als Einheit wahrnehmen. Die so transformierten Teile der Menschheit brechen dann ins All auf. Während der rückständige Rest der Menschheit in einem geheim aus dem Weltraum kontrollierten bukolischen Utopia auf Erden leben – einem »Menschenzoo«, wie es Bernal nennt –, macht sich die neue technohumane Zivilisation an die Eroberung der Weiten des Weltalls und erreicht irgendwann den Punkt der rückstandslosen Ersetzung bzw. Überwindung der menschlichen Biologie und schließlich die totale Kontrolle des Universums.

Abb. II.2

John D. Bernal (Foto nach 1964)



Quelle: Privatfotografie John Finney

Das Szenario Bernals zeigt, dass Kurzweil und andere heutige Transhumanisten in vielerlei Hinsicht lediglich Epigonen Bernals sind (dazu z. B. Kurzweil 2005). Dieser malte bereits jene radikale Veränderung der *Conditio humana* aus, die



von der heutigen transhumanistischen Bewegung als Ziel ausgegeben wird: die schrittweise Überwindung menschlicher Leiblichkeit, eine heute als »Cyber-unsterblichkeit« diskutierte ewige Existenz des individuellen Geistes, die Abschaffung menschlichen Leids und psychischer Begrenzungen sowie die Ausweitung menschlicher und schließlich nicht mehr biologisch basierter Intelligenz, die dann die Beschränkung auf den Planeten Erde überwindet und sich im Universum weithin ausbreitet (Kurzweil 2005; auch Schneider 2009, S.97). Vom heutigen Transhumanismus unterscheiden sich Bernals Visionen durch ein höheres Maß an (oder zumindest einen expliziteren) Kollektivismus. Gemeinsam haben sie mit Ersterem u. a. den quasireligiösen Überschwang sowie den Anspruch, umfassende optimistische Visionen einer Zukunft zu entwerfen, in der wesentliche Hoffnungen der traditionellen christlichen Religion im Diesseits erfüllt worden sind.

Ein weiteres grundlegendes Element der Vorgeschichte heutiger Diskussionen über neurotechnologisches HE ist jener von der Kybernetik beeinflusste, in sich sehr vielfältige Diskurs, in dem direkt auf den im Jahr 1960 öffentlich geprägten Begriff des Cyborgs Bezug genommen wird. Cyborg ist die Kurzform von »Cybernetic Organism«, also für einen sogenannten kybernetischen Organismus, und wurde von den beiden Wissenschaftlern Manfred E. Clynes und Nathan S. Kline (Clynes/Kline 1960) im Rahmen der Weltraumforschung vorgestellt. Ausgangspunkt war dabei die Überlegung, dass hinsichtlich der Anforderungen, die Aufenthalte im außerirdischen Raum an den menschlichen Körper stellen, die Veränderung menschlicher Körperfunktionen eine sinnvollere Maßnahme wäre als die, den Astronauten dort eine irdische Umgebung zur Verfügung zu stellen. Bei dieser Maßnahme könne eine Vielfalt von Mitteln zum Einsatz kommen, Ziel sei die Schaffung sich selbstregulierender Mensch-Maschine-Systeme.

Der Begriff erfuhr seitdem eine bemerkenswerte Entwicklung. Zu dieser zählt z. B. der vielbeachtete, zuerst im Jahr 1985 erschienene feministische Text »A Cyborg Manifesto: Science, Technology, and Socialist-Feminism in the Late Twentieth Century« von Donna Haraway (1991), in dem der Cyborg als Symbol für Selbstgestaltbarkeit und die Überwindung althergebrachter Identitäten gedeutet wird. Mit ihrem auch international einflussreichen Manifest hat Haraway mehrere Generationen von Gelehrten, aber auch Künstler beeinflusst, die sich in vielfältiger Weise auf Begriff und Figur des Cyborgs beziehen. Von besonderer Bedeutung ist in diesem Zusammenhang der Cyberfeminismus, der seit den frühen 1990er Jahren eine maßgebliche Rolle im feministischen Diskurs zu neuen Technologien spielt. Es ist aber darauf hinzuweisen, dass Haraway und andere einflussreiche Autorinnen in diesem Bereich – wie z. B. Katherine Hayles, die herausragende akademische Vertreterin des sogenannten Posthumanismus (Hayles 1999) – transhumanistische Visionen eines auf Leistungssteigerung abzielenden Umbaus des Körpers scharf kritisiert haben.



Eine ähnlich ambivalente Haltung zu Cyborgtechnologien und Transhumanismus findet sich in der einflussreichen Cyberpunkliteratur, einer eher düstere Zukunftsvisionen enthaltenden, in den 1980er Jahren entstandenen Spielart der Science-Fiction. In ihr wurde u. a. der Begriff »Cyberspace« geprägt – von William Gibson in »Neuromancer« (1984) –, und Ideen und Bildlichkeiten des Cyberpunk übten einen großen Einfluss auf die populäre Wahrnehmung der in diesem Bericht behandelten Technologiefelder und -visionen aus, z. B. über den bereits zuvor erwähnten Film »The Matrix«.

Aktuelle visionäre Diskurse und ihre Akteure 3.

Seit Anfang der 2000er Jahre haben Zukunftsvisionen zu allen in diesem Bericht behandelten Forschungs- und Entwicklungsbereichen in erheblichem Maße Beachtung in den Massenmedien und in verschiedenen akademischen Diskussionen gefunden. Der Transhumanismus spielt dabei als ein übergreifendes Deutungsmuster eine durchaus zentrale Rolle. In jüngster Zeit sticht in diesem Zusammenhang die Debatte über die Chancen und Risiken der KI hervor, bei der erstaunlicherweise Visionen einer sehr »starken« KI im Mittelpunkt stehen.

Die aktuellen Diskussionen über Zukunftsvisionen zu Neurotechnologien, KI und Robotik werden erheblich von in den letzten Jahren deutlich gewachsenen Akteursnetzwerken mitgeprägt, in denen u. a. Naturwissenschaftler, Ingenieure, führende Computer- und Internetindustrielle, Ethiker, Medienvertreter, Künstler und transhumanistische Aktivisten zusammenwirken. Zugleich haben die einschlägigen, auch transhumanistischen Zukunftsvisionen zunehmend Aufmerksamkeit jenseits der Massenmedien erlangt, in verschiedenen akademischen und künstlerischen Kontexten, aber auch, in geringerem Umfang, im politischen Diskurs. Im Folgenden werden zunächst einige relevante Netzwerke näher vorgestellt, um den spezifischen Charakter des zukunftsvisionären Diskurses zu Mensch-Maschine-Entgrenzungen und KI, aber auch Gemeinsamkeiten mit älteren technikvisionären Diskussionen herauszuarbeiten.⁵

5 Bis Anfang der 2010er Jahre haben diese Netzwerke vor allem in der Technikfolgenabschätzung, in der sozial- und geisteswissenschaftlichen oder philosophischen Wissenschafts- und Technikforschung sowie (in geringem Umfang) in der Forschungs- und Technologiepolitik Beachtung gefunden. Zudem fand eine Berichterstattung in Massenmedien wie der F.A.Z. statt, wobei diese Zeitung transhumanistischen und anderen stark technikfuturistischen Autoren bereits um das Jahr 2000 herum ein Forum bot (Schirmmacher 2001). Vor allem außerhalb Deutschlands waren zudem die transhumanistischen und technikfuturistischen Netzwerke Zielscheibe der Kritik einiger zivilgesellschaftlicher Organisationen (wie der ETC Group) und auch militanter, z. T. verschwörungstheoretischer Aktivisten (in Frankreich).

Themensetzung durch futuristische Netzwerke

3.1

Da in diesem Teilkapitel insbesondere auch auf den Einfluss des Transhumanismus eingegangen wird, sei eingangs noch einmal darauf hingewiesen, dass es zwar seit den 1990er Jahren eine (kleine) organisierte transhumanistische Bewegung gibt, es sich beim Transhumanismus aber vor allem um eine Weltanschauung in Form eines Ensembles von Zukunftsvisionen handelt, das bereits im ersten Drittel des 20. Jahrhunderts weitgehend entwickelt wurde (Kap. II.2) und dessen Einfluss auf das Denken über die Zukunft von Wissenschaft, Technik und Gesellschaft weit über die organisierte Bewegung hinausreicht.

Visionäre Ingenieure und die frühe transhumanistische Community

Bereits seit den 1970er Jahren bestehen z. T. überlappende Netzwerke, von denen weitreichende, vor allem transhumanistische Zukunftsvisionen zu Feldern wie der Weltraumforschung, KI, Nanotechnologie, Biotechnologie und Neurotechnologie propagiert werden. Die zentralen Akteure in diesen Netzwerken waren und sind sogenannte visionäre Ingenieure, die neben ihrer mehr oder weniger entwicklungsorientierten Forschungstätigkeit auch als Visioneers agieren, also als Konstrukteure und Propagandisten weitreichender Zukunftsvisionen (dazu z. B. McCray 2012; Schummer 2009). Ein Charakteristikum des Wirkens dieser Netzwerke ist der Umstand, dass sie mittels oft kühner Prognosen diverse Zukunftsvisionen zu den gerade genannten und zu weiteren Forschungs- und Entwicklungsfeldern miteinander verbinden (dazu und zum Folgenden Coenen 2011 sowie Heil/Coenen 2014, S. 155 f.).

Die bekanntesten Visioneers dieser Netzwerke – die bereits genannten Drexler, Kurzweil, Minsky und Moravec – gelten in ihren eigenen Forschungsfeldern und über diese hinaus entweder als »Enfants terribles« oder als besonders visionäre Denker, die zwar extreme Ansichten vertreten, dadurch aber auch die weitreichenden Potenziale der jeweiligen wissenschaftlich-technischen Entwicklungen ins Blickfeld rücken. Drexler und Kurzweil sind Schüler Minskys, und in ihren Visionen – wie auch denen Moravecs – spielt sehr »starke« KI eine zentrale Rolle. Weitere Naturwissenschaftler oder Ingenieure, die eine ähnliche Rolle wie die Genannten spielen, sind z. B. Kevin Warwick, der in der Öffentlichkeit mehr für seine Cyborgselbstexperimente und -selbstdarstellung als für seine Arbeiten in den Bereichen KI und Robotik bekannt ist, und der akademisch sehr produktive Biochemiker und Molekularbiologe George Church, der in den 2010er Jahren verstärkt mit auch über die Biologie hinausgehenden transhumanistischen Visionen hervorgetreten ist (Church/Regis 2012).



Die Edge Foundation

Eine bedeutende Rolle im zukunftsvisionären Diskurs zu Naturwissenschaft und Technik spielt die Edge Foundation mit ihrer Website edge.org. Der Verleger John Brockman hat hier ein Netzwerk von Autoren mit vor allem naturwissenschaftlichem Hintergrund geschaffen, das auch in Deutschland – u. a. aufgrund der Kooperation mit Schlüsselfiguren der hiesigen Medienlandschaft (wie z. B. Frank Schirrmacher und Hubert Burda) – meinungsbildend in Bezug auch auf Zukunftsvisionen zur KI und zu HE gewirkt und zudem auf einiges Interesse seitens führender IT-Industrieller gestoßen ist.

Seit dem Jahr 2007 findet jährlich eine Veranstaltung namens »Edge Master Class« statt, an der neben Wissenschaftlern und vereinzelt Politikern auch mehrfach Führungsfiguren der IT-Industrie (wie die Google-Gründer) teilnahmen und sich Vorträge von bekannten Wissenschaftlern (z. B. von Church und dem prominenten Biochemiker und Unternehmer Craig Venter) anhörten und mit diesen diskutierten. Weitreichende Zukunftsvisionen spielten dabei eine bedeutende Rolle. Bereits seit dem Jahr 1999 finden »The Edge Annual Dinners« statt, die auch als Milliardärsdinners bezeichnet werden. Hier haben sich – von Brockman so getaufte – »Intellektuelle der Dritten Kultur« (Naturwissenschaftler und Ingenieure, die in die Rolle des welterklärenden Intellektuellen schlüpfen, aber auch Künstler, Philosophen und andere) u. a. mit Gründern und Führungsfiguren von IT-Unternehmen wie Amazon, AOL, Google und Microsoft und anderen Vertretern der Privatwirtschaft getroffen. Der als Verleger vieler dieser »Intellektuellen der Dritten Kultur« tätige Brockman bezeichnet diese Abendessen als bemerkenswerte Treffen außergewöhnlicher Geister, von Menschen, die unsere globale Kultur umschreiben (»rewriting our global culture«). An den Dinners nahmen auch verschiedene bekannte transhumanistische Denker und Aktivisten teil wie Minsky, Kurzweil und der Philosoph Nick Bostrom, eine Schlüsselfigur des organisierten Transhumanismus, Leiter des Future of Humanity Institute in Oxford und Autor des 2014, in Übersetzung auch in Deutschland erschienenen Buchs »Superintelligence« (Bostrom 2014). Seit dem Jahr 2005 stellt die Edge Foundation zudem jährlich eine Frage und veröffentlicht die Antworten dann auf edge.org und z. T. zusätzlich in Buchform. Während fast alle Fragen allgemeiner Natur waren (z. B. »Was halten Sie für wahr, können es aber nicht beweisen?«, »Über was haben Sie Ihre Ansicht geändert?« und »Was wird alles verändern? Welche alles verändernden wissenschaftlichen Ideen und Entwicklungen erwarten Sie noch zu Ihren Lebzeiten?«), wurde im Jahr 2015 gefragt: »What do you think about machines that think?« (»Was denken Sie über denkende Maschinen?«).⁶

6 www.edge.org/annual-question/what-do-you-think-about-machines-that-think



Der organisierte Transhumanismus

Die Edge Foundation ist indes nur ein, wenn auch hervorstechendes Beispiel eines einflussreichen technikfuturistischen Netzwerkes. So hat es im Bereich der transhumanistischen Organisationen seit den 2000er Jahren einerseits Einheitsbestrebungen gegeben, andererseits Differenzierungsprozesse. Die wichtigste transhumanistische Dachorganisation – die in vielen Ländern über Mitgliedsorganisationen verfügt – nennt sich seit dem Jahr 2008 Humanity Plus (oder H+) und alte interne Gegensätze innerhalb der Bewegung sind zurückgegangen oder verschwunden. Dies betrifft vor allem den für eine Weile sehr starken Gegensatz zwischen dem älteren organisierten Transhumanismus (dem sogenannten Extropianismus mit Wurzeln in den 1970er Jahren und einer Organisationsgeschichte seit den frühen 1990er Jahren) und der u. a. von Bostrom im Jahr 1998 gegründeten World Transhumanist Association (WTA), der Vorgängerorganisation von H+. Zugleich besteht eine ganze Reihe von erklärt transhumanistischen oder transhumanistisch ausgerichteten Instituten, die erhebliche personelle Überschneidungen, aber leicht unterschiedliche thematische Ausrichtungen aufweisen. Die Mehrzahl dieser Institute hat erhebliche Förderungen vor allem von Führungsfiguren oder Unternehmen der IT-Industrie erhalten oder profitiert noch von solchen. Genannt werden können hier das britische Centre for the Study of Existential Risk der Universität Cambridge, das Future of Humanity Institute der Universität Oxford sowie sechs US-amerikanische Einrichtungen: das Future of Life Institute, das Foresight Institute, das Center for Responsible Nanotechnology, das Institute for Ethics and Emerging Technologies, das Machine Intelligence Research Institute und die auf NASA-Gelände gegründete Singularity University.

Die letztgenannte, u. a. von Kurzweil und von Google gegründete Einrichtung nimmt Bezug auf das Konzept der »Singularität« (zur Vorgeschichte des Begriffs Kap. II.2.1). Es wurde insbesondere durch Kurzweil und die sogenannten Singularitarians bekanntgemacht, eine transhumanistische Strömung, die vor allem vom Machine Intelligence Research Institute (vormals: Singularity Institute for Artificial Intelligence) ausging. In Aussicht gestellt werden u. a. die Heilung von Krankheiten, eine Welt des Überflusses und des nachhaltigen Wirtschaftens sowie schließlich, nach dem Moment der Singularität, die Verschmelzung von Mensch und Maschine, an deren Ende die völlige Überwindung des menschlichen Leibes und »kybernetische Unsterblichkeit« stehen werden.

Die NBIC-Initiative zu Converging Technologies

Vor allem in der Forschungspolitik, der Technikfolgenabschätzung und der Technikphilosophie hat die im Rahmen der Nanotechnologieförderung in den USA (Paschen et al. 2004) entstandene sogenannte NBIC-Initiative zu konvergierenden Technologien einige Beachtung gefunden (TAB 2008), wobei eine Konvergenz von Nano-, Bio-, Informations-, Kommunikations- und Neurotechnologien und den zugehörigen Wissenschaften (bzw. Nano-Bio-Info-Cogno [NBIC]) gemeint ist (Roco/Bainbridge 2003). Die Initiative wurde im Jahr 2001 von Vertretern der National Science Foundation (NSF) und des Handelsministeriums ins Leben gerufen und anfangs nicht nur von einer Reihe namhafter Persönlichkeiten aus Wissenschaft und Politik unterstützt, sondern auch durch Repräsentanten großer Unternehmen und verschiedener weiterer Institutionen der US-amerikanischen Forschungspolitik. Schlüsselfiguren und Hauptinitiatoren waren Mihail (Mike) Roco von der NSF, ein zentraler Akteur im Bereich der staatlichen Nanotechnologieförderung, und der Religionssoziologe William (Bill) Bainbridge, ein bekennender Transhumanist und in der NSF lange für die Förderung von avancierten Projekten zur Informationstechnologie zuständig.

Die NBIC-Initiative kann als ein herausragendes Beispiel für den transhumanistischen Einfluss in Forschungspolitik und Wissenschaft gelten (TAB 2008). Auch wenn sie es nicht erreicht hat, ein größeres Förderprogramm zu Converging Technologies zu inspirieren, hat sie doch (insbesondere) auf den (akademischen) Diskurs zu HE stark eingewirkt und zur neuerlichen Steigerung der Aufmerksamkeit für »starke« KI-Visionen beigetragen. Zu den im Rahmen der NBIC-Initiative entwickelten Visionen (Bainbridge 2006) gehört u. a., dass (bis zum Jahr 2025) der menschliche Körper haltbarer, langlebiger, einfacher zu reparieren und widerstandsfähiger gegen Alterungsprozesse sein wird, dass sich (bis zum Jahr 2050) menschliche Gedächtnisleistungen durch externe elektronische Speicherung steigern lassen und dass (bis zum Jahr 2070) Wissenschaftler Absichten, Überzeugungen, Sehnsüchte, Gefühle und Beweggründe als klar bestimmbare, berechenbare Prozesse verstehen und beschreiben werden. Jeder Mensch werde dann das Recht haben, die Fähigkeiten frei zu wählen, die er besitzen möchte. Schließlich wird für das Jahr 2085 vorausgesagt, dass dann Maschinen gebaut werden können, die funktional äquivalent zu einem menschlichen Gehirn sind. An der Initiative war die US-amerikanische Militärforschung beteiligt, u. a. durch die DARPA, die in dieser Zeit sehr offen HE mit diversen Mitteln sowie nichttherapeutische Kopplungen von Gehirn und Maschine als Ziele definiert hatte. Zudem hat die Rhetorik der Initiative (»Improving Human Performance«) z. T. Wurzeln in der US-Militärforschung (TAB 2008).



Zur Rolle der Computer- und Internetindustrie

Offenkundig spielen derzeit Vertreter der Computer- und Internetindustrie eine entscheidende Rolle in der Popularisierung transhumanistischer und verwandter technifikuturistischer Zukunftsvisionen. Die relevanten Netzwerke existieren z. T. – wie die Edge Foundation – bereits seit den 1990er Jahren.⁷ Zu den Schlüsselfiguren zählten bzw. zählen hier der Investor Peter Thiel – der u. a. mit der Gründung von PayPal und frühen Investitionen in Facebook ein Vermögen erzielte und diverse Transhumanisten bzw. deren Aktivitäten finanziell fördert –, Jaan Tallinn – einer der Gründer von Skype und des Centre for the Study of Existential Risk der Universität Cambridge, der zudem u. a. das Bostroms Future of Humanity Institute und das Machine Intelligence Research Institute mit Spenden unterstützte – und der 2013 verstorbene IT-Unternehmer und Sachbuchautor James Martin, dessen Spende an die Universität Oxford – die größte Einzelspende in der Geschichte der Universität – zum Aufbau der Oxford Martin School und dabei auch zur Gründung des Future of Humanity Institute und zur Förderung weiterer transhumanistischer Denker diente. Im Jahr 2006 organisierte die Oxford Martin School (zu diesem Zeitpunkt noch James Martin 21st Century School) die Konferenz »Tomorrow's People: The Challenges of Technology for Life Extension and Enhancement« (Healy/Rayner 2009) in Oxford, bei der zum ersten Mal eine große Zahl transhumanistischer Aktivisten (einschließlich Bainbridge) und transhumanistisch orientierter Denker (wie der Ethiker John Harris) in einem größeren akademischen Rahmen zusammenkamen und dabei auch auf renommierte nichttranshumanistische Wissenschaftler und Philosophen trafen, darunter bekannte Kritiker dieser Weltanschauung. Auf dieser Veranstaltung pries Martin den Transhumanismus als die Lösung für alle großen globalen Probleme. Unter den bedeutenden anwesenden Wissenschaftlern befand sich Martin Rees, Königlicher Astronom (Astronomer Royal), Kosmologe und Astrophysiker, der von 2005 bis 2010 Präsident der Royal

7 In den letzten Jahren hat in Deutschland die Medienberichterstattung zu diesen Netzwerken zugenommen und es wurden jüngst auch umfangreiche journalistische Analysen vorgelegt, die z. T. sehr kritisch oder sogar polemisch sind (Jansen 2015; Wagner 2015). Dabei stand und steht Google oft im Zentrum der Aufmerksamkeit. So heißt es in einer der sehr kritischen Analysen, dass das »transhumanistische Projekt Googles« das »größte Menschenexperiment aller Zeiten« sei, ein »Projekt der Erzeugung eines ›neuen Menschen‹, das von Transhumanisten wie Ray Kurzweil entscheidend vorgeprägt« worden sei (Jansen 2015, S.297). Das Unternehmen sei die Speerspitze des Transhumanismus. Immer wieder wird auch in Medienberichten darauf hingewiesen, dass Google beim Erwerb von KI-Unternehmen auf die Dienste Kurzweils setzt, im Bereich Robotik massive Zukäufe tätigte (u. a. den führenden Militärroboterhersteller Boston Dynamics) und im Bereich Genetik und Biotechnologie mit Calico ein Unternehmen gründete, dessen Forschungsagenda als eine Kampfansage an das Altern, vielleicht sogar den Tod gelesen werden kann.



Society war. Er spielt mittlerweile eine wichtige Rolle in der Debatte über »starke« KI und die womöglich durch sie entstehenden Risiken.

Die Singularitätsidee ist in Teilen der IT-Industrie sehr populär. Veranstaltungen des in Machine Intelligence Research Institute umbenannten Singularity Institute for Artificial Intelligence wurden von Führungsfiguren der IT-Industrie unterstützt oder als Redner besucht, z. B. von Peter Thiel und von Justin Rattner, dem ehemaligen Chief Technology Officer von Intel. Intel wiederum rief auf seinem internationalen Entwicklerforum im Jahr 2008 den Countdown auf dem Weg zur Singularität aus und führte dabei auch ein Video über Kurzweil und seine Ideen vor. Bill Gates, der Gründer von Microsoft, pries Kurzweil als wichtigsten Zukunftsdenker unserer Zeit, und zahlreiche führende Mitarbeiter von Microsoft, Intel und den neueren Internetunternehmen haben (z. T. regelmäßig) an Veranstaltungen der genannten Netzwerke oder Organisationen teilgenommen.

Der aktuelle Diskurs zu Enhancement und Cyborgs

3.2

Nachdem bereits seit längerem argumentiert wird, dass Menschen mit Herzschrittmachern, anderen Implantaten oder Prothesen per definitionem als Cyborgs gelten können, hat sich unlängst – nach Pionieraktivitäten Ende der 1990er Jahre durch den bereits erwähnten transhumanistisch gesinnten Forscher Kevin Warwick – in verschiedenen Ländern eine kleine, aber massenmedial vielbeachtete Cyborgbewegung gebildet. In Deutschland nimmt hier der in Berlin gegründete Verein Cyborg e. V. eine Vorreiterrolle ein. Zu beachten ist, dass die neuen Cyborgs keineswegs durchgängig der transhumanistischen Bewegung zuzurechnen sind. In den programmatischen Erklärungen des Cyborg e. V. zeigt sich z. B. insofern eine Distanz zum verbesserungsorientierten, HE ins Zentrum stellenden Transhumanismus, als der Verein als übergeordnetes Ziel der Körpermodifikation nicht die Steigerung der Leistungsfähigkeit oder gar die Überwindung menschlicher Körperlichkeit nennt und sich gegen jegliche »Normierung des Menschen« ausspricht (Kap. III.1.2).⁸

Der transhumanistische Diskurs zu HE ist hingegen ganz überwiegend bestimmt durch Denkfiguren der Steigerung menschlicher Leistungsfähigkeit, der verbesserten Beherrschung körperlicher Funktionen und der Emotionalität sowie letztlich – durchaus kollektivistisch im Sinne einer »Verbesserung des Men-

⁸ <http://wiki.cyborgs.cc/wiki/Manifest> (15.12.2015)



schen« gedacht – einer biologische Körperlichkeit überwindenden Weiterentwicklung des Menschengeschlechts.⁹

Die entsprechenden Debatten laufen inzwischen seit bald zwei Jahrzehnten, und einige ihrer Themen und Argumente sind, wie bereits dargelegt wurde (Kap. II.2), weit älter. Eine Vielzahl von Entwicklungen und Ideen in verschiedenen Wissenschafts- und Technikfeldern wird hinsichtlich ihrer Relevanz für Enhancement thematisiert. Dabei kommen u. a. die Bio-, Nano- und Neurotechnologien, verschiedene Bereiche der pharmazeutischen Forschung und Entwicklung, die »kosmetische« plastische Chirurgie und die Prothetik zur Sprache. Als aktuelle oder mögliche zukünftige vorrangige Einsatzbereiche von Human-Enhancement-Technologien werden u. a. das Doping im Sport, das Militär, das Bildungssystem, das sogenannte aktive Altern und die Behindernungskompensation genannt. Zahlreiche Analysen der letzten Jahre sind zu dem Schluss gekommen, dass zwischen den Visionen und der Realität gegenwärtiger Forschungs- und Entwicklungsstände insgesamt eine breite Lücke klafft (z. B. für das pharmakologische kognitive Enhancement Ferrari et al. 2012; Sauter/Gerlinger 2012). Zudem wird immer wieder darauf hingewiesen, dass eine Faszination durch Visionen einer technischen Umgestaltung des Menschen die Rationalität der politischen und gesellschaftlichen Diskurse zu Naturwissenschaft und Technik negativ beeinträchtigen könne und dass tatsächlich relevante ethische und gesellschaftliche Probleme dabei aus dem Blick geraten oder übersehen würden (Nordmann 2007; TAB 2008).

Es ist jedoch ebenfalls festzuhalten, dass im Bereich der Neurotechnologien weitreichende Enhancementvisionen zumindest zum derzeitigen Zeitpunkt weniger unrealistisch erscheinen als die Visionen eines pharmakologischen kognitiven Enhancement. Das liegt zum einen daran, dass über die Potenziale stimulierender Neurotechnologien für HE bisher weniger bekannt ist als über die medikamentöser Mittel, zum anderen eröffnen sich durch Technologien, bei denen neuronale Signale über eine technische Schnittstelle ausgelesen und in maschinelle Aktivität umgesetzt werden, zumindest für körperlich eingeschränkte, prinzipiell aber auch für alle Nutzer neue Handlungsmöglichkeiten. Auch wenn – da bei breiteren Definitionen die Debatte ihren spezifischen Sinn verlöre – die letztgenannten, »ableitenden« Technologien nur dann als HE-Technologien gelten sollten, wenn sie mit operativen Eingriffen in den menschlichen Körper einhergehen (Coenen et al. 2009), sind doch die auf sie bezogenen

9 In einer acatech-Publikation, die Innovationspotenziale der Mensch-Maschine-Interaktion behandelt, heißt es, dass der wirtschaftsliberale und libertäre Transhumanismus des Silicon Valley zwar unbestritten ein technologiefreundliches Klima schaffe, dass aber an ihm insbesondere die Erzeugung übersteigter Erwartungen an technische Lösungen gesellschaftlicher Probleme sowie ein negativer Einfluss auf den demokratischen Diskurs über neue Technologien kritisiert würde (acatech 2016, S. 58).



Zukunftsvisionen für eine Untersuchung zu Fragen der Mensch-Maschine-Entgrenzung hochrelevant.

Im Folgenden sollen Visionen zur Nutzung von Neurotechnologien für HE nicht dahingehend bewertet werden, inwieweit sie als realistische Möglichkeiten für die nähere Zukunft erscheinen (dazu aber Kap. III.1.2). Vielmehr geht es lediglich darum, einen Einblick in den stark zukunftsvisionären Diskurs zu Neurotechnologien und HE zu bieten.

Die relevanten Visionen lassen sich (nicht ganz trennscharf) folgenden Zielen zuordnen: 1. der Verbesserung der kognitiven Leistungsfähigkeit von Menschen; 2. der verbesserten Kontrolle menschlicher Emotionalität; 3. der Erweiterung der sinnlichen Wahrnehmung; 4. neuen Kommunikationsmöglichkeiten zwischen Menschen (oder auch zwischen Menschen und anderen Wesen); 5. Kopplungen von Menschen und Artefakten (Prothesen oder steuerbaren körperexternen Maschinen), bei der Letztere 6. in einem bisher unbekanntem Maß gleichsam zu Erweiterungen des menschlichen Körpers werden:

1. Visionen einer Verbesserung der kognitiven Leistungsfähigkeit von Menschen werden – vor dem Hintergrund der Annahme, dass es künftig zu einer Konkurrenz zwischen menschlicher und Maschinenintelligenz kommen werde – oft im Zusammenhang mit Visionen zu »starker« KI entwickelt. Neben älteren kybernetisch inspirierten, aus den 1950er und 1960er Jahren stammenden Visionen einer parallelen, einander bedingenden Steigerung menschlicher und künstlicher Intelligenz (z. B. Licklider 1960) werden hier allerdings in letzter Zeit – auch vor dem Hintergrund einer gewissen Desillusionierung in Bezug auf das pharmakologische kognitive Enhancement – vor allem Möglichkeiten einer Verbesserung menschlicher kognitiver Fähigkeiten mittels transkranieller Magnetstimulation oder transkranieller Gleichstromstimulation (zu diesen Technologien Kap. III.1.1 »Stimulationsverfahren«) erforscht und diskutiert (Maslen et al. 2014).
2. In letzter Zeit wird ebenfalls – insbesondere in transhumanistischen Kreisen – verstärkt diskutiert, ob die tiefe Hirnstimulation (Kap. III.1.2) die Möglichkeit eröffnen könnte, menschliche Emotionen zuverlässig zu kontrollieren und zu manipulieren. Diese Visionen stehen im Zusammenhang mit – wiederum vor allem von transhumanistischer Seite entwickelten – Visionen eines sogenannten moralischen Enhancement, bei dem menschliche Eigenschaften und Verhaltensweisen pharmakologisch oder neurotechnologisch beeinflusst werden sollen.
3. Nicht nur auf Basis des transhumanistischen Leistungssteigerungs- und Körperüberwindungsparadigmas werden bislang erfolgte, neue oder zukünftige neurotechnologische Entwicklungen (wie z. B. Cochlea- und Retina-Implantate) zum Anlass genommen, einen Zukunftsmenschen zu imaginieren, der über neue, künstliche Sinne verfügt. Diese Vision ist auch



maßgeblich für den neuen Cyborgismus (z. B. des Cyborg e. V.), der – durchaus z. T. auch in Abgrenzung vom transhumanistischen Leistungssteigerungs- und Körperüberwindungsparadigma – einen wenn nicht spielerischen, so doch experimentellen Umgang mit menschlicher Körperlichkeit propagiert, bei dem gängige Körpernormen z. B. aus feministischer Sicht hinterfragt werden.

4. Sensorische Neuroprothesen (wie Cochlea- und Retina-Implantate), aber insbesondere auch Gehirn-Maschine-Schnittstellen werden perspektivisch zuweilen als Mittel neuartiger Kommunikationsweisen zwischen Menschen (oder auch zwischen Mensch und Tier) angesehen. Von einer gewissen Relevanz ist hier auch die alte, seit Bernal bekannte Vision technotelepathischer Kommunikation. In einer Vision des Physikers und Technikvisionärs Freeman Dyson wird z. B. eine Technotelepathie mittels direkter Umwandlung von neuronalen Signalen in Funksignale und der Platzierung mikroskopischer Sender und Empfänger im Gehirn ausgemalt.¹⁰
5. Die wohl populärsten Visionen einer Cyborgisierung sind auf BMI-Technologien gerichtet, von denen erwartet wird, dass Mensch und Maschine umfassend und tiefgreifend miteinander verschmelzen und dass künstliche Körperteile, insbesondere Gliedmaßen, eine neue technobiologische Körperlichkeit ermöglichen werden. So wurde z. B. im Rahmen der NBIC-Initiative für die 2040er Jahre prognostiziert, dass Soldaten die Fähigkeit haben werden, nur durch das Denken von Befehlen – oder sogar vor Formulierung des Befehls in ihrem Geist – Fahrzeuge sowie Waffen- und andere Kampfsysteme ohne jegliche Zeitverzögerung zu kontrollieren (Bainbridge 2006). Jüngst hat zudem eine Studie (Herff et al. 2015) die Erwartung geweckt, dass in Zukunft eine weitgehende, ganze Sätze rekonstruierende Spracherkennung aus Gehirnströmen durch neue, sogenannte Brain-to-Text-Technologien möglich sein wird. In Bezug auf Gliedmaßenprothesen wird immer wieder – und nicht nur von transhumanistischer Seite – die Erwartung geäußert, dass das allgemein akzeptierte Ziel, funktional und sensorisch natürlichen Gliedmaßen entsprechende Prothesen zu entwickeln, hin zu einer – eventuell auch militärisch nutzbaren – Cyborgisierung erweitert werden könnte, bei der dann zusätzliche Funktionen oder sogar zusätzliche Gliedmaßen bereitgestellt würden.
6. Ungeachtet des teils zweifelhaften Realitätsgehalts dieser Visionen hat es den Anschein, dass einige der in den HE- und Cyborgdebatten thematisierten Technologien einen massiven Wandel im Mensch-Technik-Verhältnis anzeigen. So dürfte die Direktkopplung von komplexeren intellektuellen Leistungen mit Computertechnik prinzipiell machbar sein (Herff et al.

¹⁰ http://edge.org/q2009/q09_3.html#dyson (15.12.2015)

2015), und auch dokumentierte emotionale und andere Auswirkungen der Nutzung von stimulierenden Neurotechnologien (z. B. Dubiel 2006; Kap. III.1.2) geben Anlass dazu, die Möglichkeit einschneidender Veränderungen im Mensch-Maschine-Verhältnis zu diskutieren.

Gerade auch vor dem Hintergrund dieser möglichen Anzeichen für einen signifikanten Wandel im Mensch-Technik-Verhältnis ist der Umstand bemerkenswert, dass sich Visionen und erste Praktiken der Cyborgisierung nicht allein im transhumanistischen Rahmen entfalten. Während in diesem Rahmen – und in ähnlichen technikfuturistischen Kontexten – relativ umstandslos eine Verbesserungs-, Körperbeherrschungs- und Entleibungslogik zugrunde gelegt wird, die angesichts der vielfältigen Aspekte und Aussichten der technischen Entwicklung unterkomplex erscheint, werden zunehmend sowohl in Kunstprojekten und anderen kulturellen Aktivitäten als auch seitens der (ja keineswegs ausschließlich transhumanistischen) Cyborgismusbewegung neue Verständnisse des Mensch-Technik-Verhältnisses wie auch neue Konzepte der Identität im technowissenschaftlichen Zeitalter entwickelt (oder zumindest als wünschenswert angesehen). Ohne dies in seiner Bedeutung überschätzen zu wollen, lässt sich vermuten, dass die – wahrscheinlich auch durch einen Informationsvorsprung verursachte – Dominanz transhumanistischer Perspektiven im Diskurs zukünftig abnehmen wird.

Der aktuelle Diskurs zu einer Superintelligenz

3.3

Angesichts der erheblichen Kritik, die Anhänger der »starken« KI wegen ihres Anfachens zukunftsvisionärer Diskussionen auf sich gezogen haben (Kap. II.2), ist es durchaus überraschend, dass die Zeit seit dem Jahr 2000 durch einen erheblichen neuerlichen Bedeutungszuwachs »starker« KI-Visionen gekennzeichnet ist. Bemerkenswert ist, dass diese Renaissance des Themas »starke« KI weniger innerhalb der relevanten Forschungs- und Entwicklungsbereiche selbst stattgefunden hat als vielmehr in übergreifenden Debatten über die Potenziale vieler verschiedener naturwissenschaftlich-technischer Felder (z. B. im Rahmen der NBIC-Initiative). Im Verlauf der 2010er Jahre setzte dann eine Debatte über eine »starke« KI ein, die im erheblichen Maße in den Massenmedien ausgetragen wird und zu der – neben (oder in) den seit den 1990er Jahren gewachsenen transhumanistischen Netzwerken – u. a. auch eine ganze Reihe bekannter Wissenschaftler (wie z. B. Stephen Hawking) und Führungsfiguren der US-amerikanischen Computer- und Internetindustrie (wie z. B. Elon Musk) beiträgt. Diese Diskussion zeigt exemplarisch die hohe Themensetzungs- und Mobilisierungsfähigkeit der in Kapitel II.3.1 dargestellten Netzwerke. In den Medien, aber auch auf öffentlichen Diskussionsveranstaltungen, die von wissenschaftlichen Ein-



richtungen organisiert werden, wird dabei z. T. mit plakativen Formulierungen gearbeitet, die das Kernthema der Debatte auf den Punkt bringen (»Mensch, Maschine! Wer wird künftig den Ton angeben?«; »Bald sind Maschinen intelligenter als wir. Werden sie ... Freund oder Feind?«).

Bemerkenswert ist der Umstand, dass sich die in der aktuellen Debatte vorherrschenden Denkfiguren, Zukunftsvisionen und Argumente bereits seit längerer Zeit in transhumanistischen Zirkeln verbreiten und weitgehend identisch mit den in Kapitel II.2.1 erwähnten Spekulationen I. J. Goods zu einer »Intelligenzexplosion« sowie daran anknüpfenden Überlegungen Vernor Vinges und anderer zur sogenannten Singularität sind. Erwähnt werden können hier z. B. die populärwissenschaftlichen Arbeiten Hans Moravecs seit den 1980er Jahren (z. B. Moravec 2001) sowie neuere Wortmeldungen aus dem Machine Intelligence Research Institute (vormals Singularity Institute for Artificial Intelligence) und von Nick Bostrom. Letzterer antwortete z. B. auf die Edge-Frage des Jahres 2009 »What will change everything?« (»Was wird alles verändern? Welche alles verändernden wissenschaftlichen Ideen und Entwicklungen erwarten Sie noch zu Ihren Lebzeiten?«), dass dies die Heraufkunft einer »Superintelligenz« sein werde.¹¹ Weitgehend auf Überlegungen Goods, Vinges' und Moravecs basierend, plädiert Bostrom dafür, dass wir uns schon jetzt auf eine Intelligenzexplosion vorbereiten sollten, bettet die Vision einer »starken« Super-KI in die transhumanistische Großzählung ein und warnt, dass die Entstehung einer dem Menschen nicht freundlich gesonnenen Super-KI ein für die Menschheit existenzielles Risiko bedeuten könne (siehe auch Bostrom 2014).

Die Debatte gewann an Schwung und verließ die transhumanistisch-technikfuturistischen Zirkel, nachdem sich im Frühjahr 2014 der weltweit auch einer breiteren Öffentlichkeit bekannte theoretische Physiker und Sachbuchautor Stephen Hawking, der Computerwissenschaftler und KI-Forscher Stuart Russell, der Kosmologe Max Tegmark (der einer der Gründer des Future of Life Institute ist und Spenden von Elon Musk für die Erforschung existenzieller Risiken durch fortgeschrittene KI erhielt) sowie Frank Wilczek, Nobelpreisträger für Physik des Jahres 2004, zu Wort gemeldet hatten. In einem Artikel für die »Huffington Post« (19.4.2014)¹² und die Zeitung »The Independent« (1.5.2014) sprechen sich die Autoren gegen eine denkfaule Selbstgefälligkeit in Bezug auf mögliche superintelligente Maschinen aus.¹³ Die vier Autoren beto-

11 <https://edge.org/response-detail/10228> (15.12.2015)

12 www.huffingtonpost.com/stephen-hawking/artificial-intelligence_b_5174265.html (15.12.2015)

13 Der Artikel erschien in der Woche der Kinopremiere des bereits erwähnten Films »Transcendence«, in dem ein transhumanistisch gesinnter Wissenschaftler, der nach einem Attentat durch militante KI-Gegner mittels Mind Uploading weiterexistiert, durch das Internet global die Kontrolle zu gewinnen versucht. Im Film sieht man den IT-Unternehmer Elon Musk eine Rede des Protagonisten im Stil Kurzweils lauschen.



nen, dass es sich als der Menschheit größter Fehler herausstellen könne, wenn hochintelligente Maschinen als bloße Science-Fiction abgetan werden. Ein Computer, der eine Quizshow im Fernsehen gewinnt – gemeint ist die »Watson« getaufte KI von IBM, die mittlerweile u. a. im Gesundheitswesen, bei der Analyse sozialer Netze und von Banken und Versicherungen (auch in Deutschland) eingesetzt wird (z. B. Fromme 2015) –, selbstfahrende Autos und digitale persönliche Assistenten wie »Siri« seien lediglich Symptome eines informationstechnologischen Rüstungswettlaufs, der durch nie dagewesene Investitionen und eine sich ständig verbessernde theoretische Grundlage ermöglicht werde. Blicke man weiter in die Zukunft, ließen sich keine grundsätzlichen Grenzen des Möglichen ausmachen, und ein explosionsartiger Übergang in ein neues Zeitalter, in dem KI Wirtschaft, Wissenschaft, Militärwesen und Politik bestimmen wird, sei durchaus denkbar. Während die kurzfristigen Auswirkungen der KI davon abhingen, wer sie kontrolliert, stelle sich längerfristig die Frage, ob sie sich überhaupt kontrollieren lässt. Trotz dieser Aussichten gebe es wenige ernsthafte Forschung zum Thema. Nur kleine gemeinnützige Einrichtungen wie das Centre for the Study of Existential Risk der Universität Cambridge, das Future of Humanity Institute, das Machine Intelligence Research Institute und das Future of Life Institute stellten sich dieser Aufgabe.

Auf diesen Artikel folgte dann eine ganze Reihe weiterer Wortmeldungen ähnlichen Tenors, u. a. von den Industriellen Bill Gates (Microsoft), Elon Musk und Peter Thiel sowie noch einmal von Hawking, der – wie u. a. auch Bostrom, Church, Musk, Rees, Russell und Wilczek – zum wissenschaftlichen Beirat des von Tallin, Tegmark und anderen gegründeten Future of Life Institute gehört (Davies 2014).¹⁴ Gates bekundete, sich ebenfalls Sorgen hinsichtlich einer zukünftigen Superintelligenz zu machen und Hawking stellte seine Aussagen zu KI in den Kontext einer umfassenderen transhumanistischen Vision, indem er Überlegungen zu einem gentechnischen HE und zur menschlichen Kolonisation des Weltraums anstellte (Davies 2014). Der Bekanntheit und dem professionellen Hintergrund der Mahner entsprechend kam es zudem zu einer Welle von Beiträgen in Massenmedien.

Die Edge-Frage des Jahres 2015 »What do you think about machines that think?« (»Was denken Sie über denkende Maschinen?«) wurde von knapp 200 Personen beantwortet. Neben Wortmeldungen im Tenor der aktuellen Diskussion, bei denen auch von der Möglichkeit einer selbstbewussten KI ausgegangen wird, finden sich zahlreiche Beiträge, in denen die Sorge um eine men-

14 Ähnlich ist die Situation hinsichtlich des Centre for the Study of Existential Risk der Universität Cambridge; zu dessen Gründern zählen Rees und Tallin und unter dessen wissenschaftlichen Beratern gehören Bostrom, Church, Hawking, Musk, Russell und Tegmark.



schenfeindliche Superintelligenz als unbegründet abgetan, aber davor gewarnt wird, zu viel Autorität an unintelligente Maschinen abzutreten, also Ergebnisse purer Rechenkraft zur nicht mehr hinterfragten Grundlage von Entscheidungen zu machen. Zudem wird oft dafür plädiert, statt über Mensch-gegen-Maschine-Visionen lieber über ökonomische oder politische Aspekte der KI-Entwicklung und -Nutzung nachzudenken und zu diskutieren.¹⁵

Im Dezember 2015 gab dann Musk bekannt, dass er zusammen mit Mitstreitern vor allem aus dem Bereich der Computer- und Internetindustrie (u. a. Thiel) eine »Open AI« (»offene KI«) benannte Einrichtung fördern wird, in der renommierte Wissenschaftler und Ingenieure KI-Forschung betreiben und die Ergebnisse öffentlich verfügbar machen sollen. Diese Positionierung wurde allerdings auch als bloße Verschleierung des Umstands gewertet, dass der Mahner Musk nun im KI-Bereich selbst die Konkurrenz mit Google und anderen von ihm als rein profitorientiert dargestellten Firmen aufnimmt (Tanriverdi 2015).

Fazit

4.

Im Wesentlichen entwickeln sich die visionären Diskussionen über die Zukunftsaussichten »starker« KI und über neurotechnologiebasierte Mensch-Maschine-Entgrenzungen nach einem Muster, das aus älteren zukunftsvisionären Technikdebatten – z. B. über Nanotechnologie und Converging Technologies – bekannt ist: In mehr oder weniger eng miteinander verbundenen Zirkeln prominenter Naturwissenschaftler, Ingenieure, Autoren und Industrieller werden z. T. sehr weitreichende Zukunftserwartungen diskutiert, die stark durch die transhumanistische Weltanschauung geprägt sind. Diese Erwartungen sind weitgehend deckungsgleich mit Spekulationen, die bereits vor vie-

15 Auch in Deutschland haben die Beiträge zur neuen Debatte über KI mittlerweile zugenommen. Beispielfähig nennen lässt sich hier ein Diskussionspapier zu Chancen und Risiken der KI, an dem u. a. mit dem Philosophen Thomas Metzinger ein im Edgework Netzwerk aktiver Forscher beteiligt war (Mannino et al. 2015). Auch in diesem wird die Erwartung geäußert, dass eine Intelligenzexplosion nach dem von Good vorhergesagten Muster stattfinden wird. Beispiellose ethische Herausforderungen würden sich daraus ergeben. Auch die Entwicklung phänomenaler Zustände in künstlichen Intelligenzen sei für die Zukunft nicht auszuschließen, sie könnten dann über (Selbst-)Bewusstsein und insbesondere auch über subjektive Präferenzen und Leidensfähigkeit verfügen. In der Studie wird sogar, was verwundert, von einem »soliden Konsens« darüber gesprochen, dass ein künstliches Bewusstsein prinzipiell möglich ist. Das Thema KI-Sicherheit verdiene auf jeden Fall in Politik und Forschung stärkere Beachtung. Es gelte, institutionell sicherheitsfördernde Maßnahmen auszuarbeiten, z. B. durch die Vergabe von Forschungsgeldern für entsprechende Projekte. Entwicklungsprojekte zu gehirnanalogen KI-Architekturen sollten unter die Aufsicht von Ethikkommissionen gestellt werden.

len Jahrzehnten von Wissenschaftlern wie J. D. Bernal und I. J. Good angestellt wurden. Auch dieser epigonale Zug der aktuellen Diskussionen ist aus anderen neueren Technikdebatten bekannt. Nur wenig überspitzt lässt sich sagen, dass z. B. im Fall der aktuellen Debatte über Superintelligenz (bzw. eine sehr »starke« KI) eine relativ kleine Gruppe von bekannten Wissenschaftlern, Autoren und Industriellen – die selbst z. T. Institute gegründet haben, die sich mit den diskutierten Zukunftsaussichten befassen – für das Thema in der massenmedialen Öffentlichkeit Aufmerksamkeit geschaffen haben, wobei sie sich im Wesentlichen auf Spekulationen aus den 1960er Jahren stützen, ergänzt durch Verweise auf neuere technische Entwicklungen, die allerdings mit Visionen einer sehr »starken« KI nur wenig zu tun haben.

Neu – z. B. im Vergleich zum Diskurs über Nanotechnologie – ist aber, dass führende Computer- und Internetindustrielle selbst die Debatte inhaltlich mitprägen. Zwar spielen hier wieder z. T. dieselben Netzwerke (wie z. B. die Edge Foundation Inc.) eine zentrale Rolle, die schon um das Jahr 2000 herum die Diskussionen über Nanotechnologie, KI und HE erheblich beeinflusst haben, und auch schon in diesen Diskussionen kam es – insbesondere durch Bill Joy (Schirmacher 2001) – zu einflussreichen Wortmeldungen aus der Computer- und Internetindustrie. Inzwischen hat sich aber die Zahl der direkten Beiträge von Industriellen zur Diskussion deutlich erhöht, wie auch die Zahl der Führungsfiguren aus der Computer- und Internetindustrie, die Forschung zu technikvisionären Themen und transhumanistische Aktivisten finanziell fördern. Überdies zeigen die Akquiseaktivitäten und Kooperationen von Firmen wie Google bzw. der Holding Alphabet, dass in der Industrie ein signifikantes Interesse an der Entwicklung avancierter KI- und Robotersysteme besteht, und einige öffentliche Äußerungen von Führungsfiguren der Computer- und Internetindustrie legen die Vermutung nahe, dass die Charakterisierung des Transhumanismus als einer Art Ersatzreligion nicht gänzlich aus der Luft gegriffen ist.

Auch wenn z. B. verschiedentlich die beiden Google-Gründer weitreichende Visionen zu KI und zur leistungssteigernden Verschmelzung von Mensch und Maschine öffentlich geäußert haben, ist es aber eine unzulässige Verkürzung, wenn das Unternehmen als Speerspitze speziell des Transhumanismus bezeichnet wird. Gleiches gilt allgemein für eine Reduktion des Technikfuturismus in der Computer- und Internetindustrie auf den Transhumanismus. Es zeigt sich hier vielmehr ein umfassender, gegenwartsbezogener Fortschrittsoptimismus, hinter dem die weiterreichenden Zukunftsvisionen des Transhumanismus stehen können, aber nicht müssen.

In diesem Zusammenhang ist ein SPIEGEL-Interview mit Larry Page, einem der Google-Gründer aufschlussreich (Schulz 2015), das dazu dienen kann, die nichttranshumanistischen Teile des rationalen Kerns zukunftsvisionärer Diskussionen zu KI und anderen Themen herauszuarbeiten. Unter Verweis u. a.



auf den Fortschrittsoptimismus früherer Weltausstellungen kritisiert Page ein angeblich in Europa stark verbreitetes Weltbild, das durch Furcht vor Technik gekennzeichnet sei. Angesichts des von ihm wahrgenommenen beständigen technischen und gesellschaftlichen Fortschritts verstehe er diese Technikfurcht nicht, sie komme ihm irrational vor und ergebe wissenschaftlich keinen Sinn. Auf die Frage des Interviewers, ob die großen Visionen letztlich nicht einfach nur dazu dienten, ganz banale Geschäfte schönzufärben, entgegnet Page, dass sich große Ideen am schnellsten auf globaler Ebene durchsetzen ließen, wenn man sie mit einem Produkt bzw. einem Geschäft verbindet.

Diese Art eines z. T. altmodisch anmutenden Fortschrittsoptimismus stellt den Rahmen dar, innerhalb dessen sich die stärker technikfuturistischen und auch die transhumanistischen Zukunftsvisionen mit Unterstützung von Teilen der Computer- und Internetindustrie entfalten können. Befeuert werden die zukunftsvisionären Diskussionen von immer neuen Erfolgsmeldungen aus den Firmen und Universitäten, wie z. B. im Dezember 2015 die Meldung aus dem von Google und der NASA aufgebauten Quantum Artificial Intelligence Lab, dass dort ein funktionsfähiger Quantencomputer und damit potenziell eine für die KI-Entwicklung zentrale Technologie entwickelt worden sei. Dabei bleibt aber oft unklar, ob die in den Diskussionen angeführten realen Fortschritte für die jeweiligen Visionen auch tatsächlich von Bedeutung sind.

Abschließend bleibt zunächst Folgendes festzuhalten: Einerseits sollte der Umstand, dass ein relativ kleiner Kreis von – sich einander in ihren Spekulationen und sonstigen Überlegungen bestätigenden, z. T. organisatorisch miteinander vernetzten – Akteuren die zukunftsvisionären Diskussionen wesentlich prägt, nicht zum Anlass genommen werden, die Relevanz dieser Diskussionen über Mensch-Maschine-Entgrenzungen zu ignorieren. Andererseits sollten aber der erhebliche intellektuelle Einfluss oder die ökonomische Macht vieler Beteiligter nicht dazu verleiten, ohne nähere Prüfung eine besondere Dringlichkeit oder Bedeutung der verhandelten Themen vorauszusetzen. Letztlich haben wir es hier fast durchgängig mit z. T. schon sehr alten Zukunftsvisionen zu tun, deren tatsächliche Basis angesichts des aktuellen Forschungs- und Entwicklungsstandes aber immer wieder weiterer Klärung bedarf. Solange keine neuen naturwissenschaftlich-technischen oder medizinischen Durchbrüche erzielt werden, die »starke« KI-Visionen oder Hoffnungen auf ein signifikantes HE rechtfertigen, hat die Befassung mit diesen Visionen eher den Charakter von Gedankenspielen oder kultur- und geisteswissenschaftlichen Überlegungen.

Neben Analysen, in denen die Zukunftsvisionen an der Realität gemessen werden, erscheinen auch Anstrengungen angezeigt, die Mechanismen und Dynamiken technikfuturistisch geprägter Diskussionen besser zu verstehen. Seite an Seite mit Welterlösungsvorstellungen finden sich in diesem »technikvisionären Marketing« (Coenen 2009) i. d. R. auch immer Visionen eines Endes der



Menschheit oder ähnlicher Katastrophen, und beides dient zur Schaffung von Aufmerksamkeit für noch nichtexistierende Technologien. Neben der Aufmerksamkeit für die Argumentationsweisen und Themensetzungsfähigkeit des Technikfuturismus ließen sich zudem auch verstärkt Anstrengungen unternehmen, den zukunftsvisionären Diskurs über Technologien der Mensch-Maschine-Entgrenzung (wieder) offener zu gestalten. Einseitig auf eine Zukunftsvision zu fokussieren – ganz gleich, ob diese dann als wünschenswert oder als albtraumhaft dargestellt wird –, birgt u. a. die Gefahr in sich, auf tatsächliche Herausforderungen der Zukunft nicht ausreichend vorbereitet zu sein.





Neurotechnologien**1.**

Das menschliche Gehirn gehört zweifelsohne zu den komplexesten Strukturen im bislang bekannten Universum. Allein schon die schiere Anzahl von schätzungsweise bis zu 100 Milliarden Neuronen – jener Gehirnzellen, die als Kernbaustein der Reiz- und Signalverarbeitung fungieren – ist beeindruckend. Doch die eigentliche Informationsverarbeitung und damit die außerordentlichen Fähigkeiten des gesamten Nervensystems basieren auf den mannigfachen Verknüpfungen seiner neuronalen Untereinheiten (dazu und zum Folgenden Hofmann 2014, S.9). Ein einzelnes Neuron kann Eingangsinformationen von bis zu tausend anderen Neuronen erhalten, sodass die Anzahl der synaptischen Verschaltungen mit 100 Billionen die Anzahl der Sterne in der Milchstraße um drei Größenordnungen übertrifft. Hinzu kommt, dass diese Verknüpfungsmuster nicht statisch sind, sondern plastisch, d.h. sich in Abhängigkeit von eingehenden Reizen äußerst dynamisch verändern.

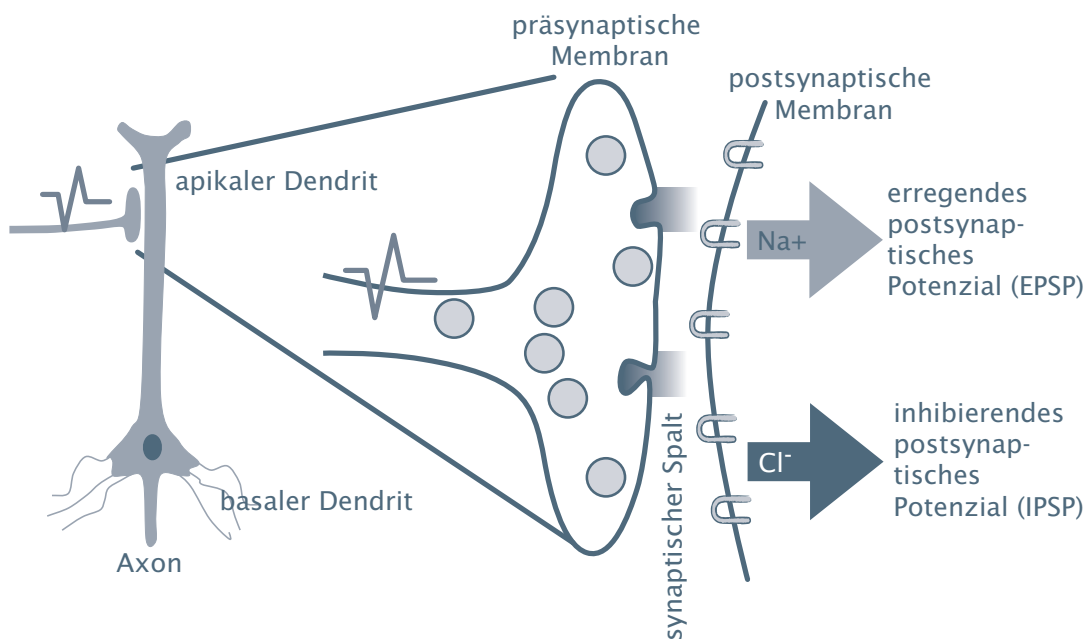
Seit vielen Jahrzehnten bemühen sich die Neurowissenschaften, die außerordentliche Komplexität des Gehirns experimentell zu entschlüsseln. Besonders in den letzten Jahren konnten – zum einen dank immer ausgefeilterer Methoden wie bildgebender Verfahren, zum anderen dank Fortschritten in verwandten Grundlagenbereichen (insbesondere der Molekularbiologie und der Informatik) – zahlreiche Erkenntnisse gewonnen werden, die nicht nur neue Einsichten, sondern zunehmend auch technische Eingriffe in das menschliche Gehirn ermöglichen (Hennen et al. 2008). Diese Entwicklung birgt große medizinische Potenziale, aber mindestens ebenso große ethische Sprengkraft. Denn das Gehirn als »Zentralorgan des Menschen« (Clausen 2010b, S.1) ist in die Regulierung zahlreicher körperlicher, motorischer sowie psychischer Funktionen involviert und damit Ausgangspunkt sowohl für die Herausbildung von (Selbst-) Bewusstsein wie auch, bei Schädigungen bzw. Funktionsstörungen, für zahlreiche Erkrankungen neurologischer (Schlaganfall, Demenzen, Parkinsonkrankheit, Epilepsie etc.) und psychiatrischer Art (Depression, Zwangs- und Angststörungen etc.). Die Entwicklung von Neurotechnologien weckt somit nicht nur Hoffnungen auf ganz neue medizinische Optionen zur Behandlung dieser Krankheiten oder zur Kompensation ihrer Folgen, sondern lässt außerdem Befürchtungen wach werden, dass zukünftig auch das gesunde Gehirn (und mit ihm das menschliche Selbst) zunehmend dem technologischen Gestaltungswillen preisgegeben werden könnten. Inwiefern diese Hoffnungen oder auch Ängste begründet sind, soll im Folgenden anhand eines fundierten Überblicks über den aktuellen Stand von Forschung und Entwicklung aufgezeigt werden

**Wissenschaftliche Grundlagen:
neuroelektrische Schnittstellen**

1.1

Trotz nach wie vor gravierender Wissenslücken, insbesondere was übergeordnete Hirnfunktionen anbelangt, ist das »Ein-Mal-Eins des Gehirns« (Monyer et al. 2004) auf zellulärer und molekularer Ebene schon recht gut entschlüsselt. So weiß man bereits seit einigen Jahrzehnten, dass die Signalverarbeitung im Gehirn auf einfachen elektrophysiologischen Grundprinzipien beruht. Die grundlegende Einheit neuronaler Kommunikation bilden demnach elektrische Potentialschwankungen, die sogenannten Aktionspotenziale, die durch das zeitweilige Öffnen und Schließen von Ionenkanälen zustande kommen und sich vom Zellkern aus in Richtung der Synapsen ausbreiten (dazu und zum Folgenden Hofmann 2014, S. 12 f.).

Abb. III.1 Schematische Darstellung eines ankommenden Aktionspotenzials



Quelle: nach Hofmann 2014, S. 19

Sie bewirken auf ihrer terminalen Seite die Ausschüttung der zelleigenen Botenstoffe (Neurotransmitter) in den synaptischen Spalt und sind ihrerseits eine natürliche Reaktion des Neurons auf die eingehenden synaptischen Signale anderer Neurone (Abb. III.1). Die Tatsache, dass die Reiz- und Signalverarbeitung im Nervensystem primär auf elektrischem Wege erfolgt, hat vielfältige Möglichkeiten eröffnet, mit neuronaler Aktivität technologisch zu interagieren. Dabei kann zwischen den Richtungen der Interaktion differenziert werden, je nach-

dem, ob Hirnsignale aus dem Nervensystem abgeleitet werden (ableitende Messverfahren) oder die Funktionsweise des Gehirns bzw. Nervensystems durch Zuführung elektrischer Ströme beeinflusst wird (Stimulationsverfahren).

Ableitungsverfahren

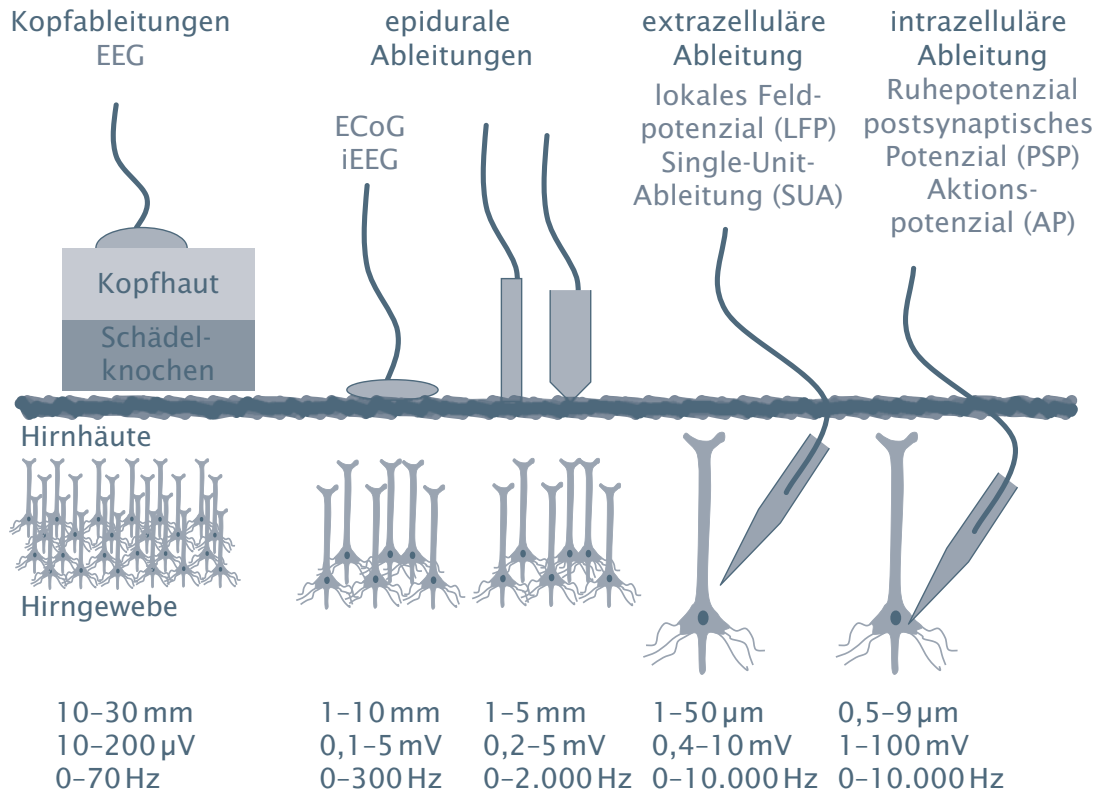
Der Goldstandard, um Aktionspotenziale und damit neuronale Aktivität zu messen, stellt die *intrazelluläre Ableitung* dar (dazu und zum Folgenden Hofmann 2014, S.16 u. 18). Dabei wird eine Mikroelektrode direkt in das Zytoplasma einer Nervenzelle eingeführt, eine zweite Elektrode in ein neutrales Medium außerhalb. Dieses Verfahren hat den Vorteil, dass sich Potenzialänderungen in einzelnen Zellen sehr präzise erfassen lassen – da ein Aktionspotenzial etwa 100 mV groß ist, waren die hierfür erforderlichen Anforderungen an Verstärkertechnik und Darstellung bereits in den 1940er Jahren erfüllt. Gleichwohl ist das Verfahren schwierig zu bewerkstelligen, da es in der Tiefe des Gehirns eine große Herausforderung darstellt, eine einzelne Zelle erstens zu lokalisieren und zweitens deren Zellmembran punktgenau zu durchstoßen, ohne die Funktionsfähigkeit der Zelle zu beeinträchtigen. Aus diesem Grund kommen zur Datengewinnung aus dem Gehirn häufig einfacher umsetzbare Messverfahren zur Anwendung, wie extrazelluläre, epidurale oder Kopfableitungen (Abb. III.2). Man zahlt zwar für die einfachere Methodik der extrazellulären Ableitung mit einer geringeren Präzision sowohl bei Signalqualität als auch Zielgenauigkeit – eine gewünschte Zelle lässt sich nicht mehr gezielt kontaktieren –, gewinnt dafür aber eine große Freiheit bei der Anzahl der Ableitpunkte im oder am Gewebe.

Bei *extrazellulären Ableitungen* werden Tiefenelektroden in die Nähe aktiver Neurone eingeführt (dazu und zum Folgenden Hofmann 2014, S.45 ff.). Damit lassen sich zwar Aktionspotenziale nicht direkt messen, dafür aber die indirekten Potentialschwankungen im extrazellulären Medium. Um so die Aktivität einzelner Neurone abbilden zu können (sogenannte Einzelzellableitungen, wie sie für elabourierte Brain-Machine-Interfaces oder aus medizinischen Gründen erforderlich sein können), ist nicht nur eine präzise Platzierung nahe der Zelle,¹⁶ sondern vor allem auch der Einsatz von Mikroelektroden erforderlich, die ungefähr die gleiche Größe wie die ableitenden Zellen haben (ca. 0,02 mm Durchmesser). Mit herkömmlichen Makroelektroden (1,2 mm Durchmesser) sind hingegen nur sehr unspezifische, räumlich schlecht aufgelöste Messungen möglich, da aufgrund der großen Elektrodenfläche alle elektrischen Signale der Umgebung erfasst werden, die von mehreren Millionen Zellen stammen können. Damit

16 Auch mit geeigneten Elektroden lassen sich neuronale Aktionspotenziale im Gehirn nur bis zu Entfernungen von 50 bis 100 µm extrazellulär feststellen (Hofmann 2014, S.20).

derartige extrazelluläre Potenziale, die im Mikrovoltbereich beginnen, sichtbar gemacht werden können, müssen die Signale um den Faktor 100.000 verstärkt werden, was moderne Verstärkertechnik erforderlich macht.

Abb. III.2 Illustration von ableitenden Verfahren



Quelle: nach Hofmann 2014, S. 21

Sowohl bei extra- als auch bei intrazellulären Tiefenableitungen handelt es sich um mikroskopische Messverfahren, die räumlich hochspezifische Messungen ermöglichen. Zur Erfassung der Hirnaktivität als Ganzes sind sie dementsprechend nicht geeignet. Hierzu ist man gezwungen, auf makroskopische Verfahren zurückzugreifen, die bei der Hirnrinde (Elektrokortikogramm) oder der Kopfoberfläche (Elektroenzephalografie) ansetzen. Dabei macht man sich das Prinzip zunutze, dass Ströme im Schädelinneren relativ gut geleitet werden, sodass durch die synchrone Aktivität großer Neuronengruppen eine nennenswerte Stromstärke entsteht, die auch auf der Hirn- bzw. der Kopfhaut nachgewiesen werden kann (Hofmann 2014, S.22). Zu beachten ist allerdings, dass Verfahren der Oberflächenableitung nur eine sehr beschränkte Tiefenschärfe aufweisen und im Wesentlichen nur die Aktivität der Großhirnrinde abbilden können. Da sich dort der Sitz der meisten kognitiven Funktionen befindet, lassen sich aus



den gewonnenen Daten dennoch interessante klinisch-neurologische Einblicke gewinnen – vorausgesetzt, es gelingt, die relevanten Signale aus dem starken Hintergrundrauschen herauszufiltern. In dieser Hinsicht gibt es zwischen den beiden Verfahren prinzipielle Unterschiede, die mit dem unterschiedlichen Grad an Invasivität zu tun haben.

Bei *epiduralen Oberflächenableitungen*, auch Elektrokortikogramm genannt (ECoG), wird ein Elektrodengitter (grid) aus Platin-Iridium- oder Edelmetallelektroden unmittelbar auf (epidural) oder unter (subdural) der harten Hirnhaut angebracht (dazu und zum Folgenden Hofmann 2014, S. 43 ff.). Dazu wird dem lokalanästhesierten und wachen Patienten ein Teil der Schädeldecke entfernt und steril zwischengelagert, um anschließend mittels eines aufgelegten Elektrodengitters elektrische Signale direkt von der Hirnhaut aufzuzeichnen. Auf diese Weise lässt sich die Aktivität der oberflächlichen Großhirnrinde nah an der Quelle erfassen, was räumlich recht gut aufgelöste Ableitungen ermöglicht – im Prinzip ist es möglich, wie Forschungen gezeigt haben, aus diesen Daten die Bewegungsprofile einzelner Körperteile zu ermitteln (Shih et al. 2012, S. 272). Aufgrund ihrer guten räumlichen Auflösung stellen Elektrokortikogramme seit den 1950er Jahren eine wichtige Methode dar, um die Quelle – den epileptogenen Herd oder Fokus – medikamentös nicht behandelbarer Epilepsie zu lokalisieren. Da natürlich nicht gewährleistet ist, dass während der wirkenden Lokalanästhesie ein Anfall vorkommt, werden heutzutage Epilepsiepatienten in einer ersten Operation mit dem Elektrodengrid ausgestattet und der Schädel wieder verschlossen. Diese Patienten verbleiben dann für bis zu drei Wochen auf speziell ausgestatteten Epilepsiestationen, wo sowohl 24 Stunden am Tag das Elektrodengrid abgeleitet wird als auch eine konstante Video- und Vitalzeichenüberwachung gewährleistet ist. Die Bereitschaft der Patienten vorausgesetzt können währenddessen auch einfache neurowissenschaftliche Untersuchungen unter kontrollierten Bedingungen durchgeführt werden.

Als nichtinvasives Verfahren ist die *Elektroenzephalografie* (EEG) ohne komplizierte Operation und damit völlig risikolos durchführbar. Elektrisch leitende Metallnäpfchen (Gold, Edelstahl oder Platin) oder Keramikringe (Silber bzw. Silberchlorid) werden hier an vorgegebenen, standardisierten Positionen mittels einer Kopfhaube oder eines flexiblen Netzes direkt auf der Kopfhaut fixiert.¹⁷ Durch die größere Distanz zur Quelle ergibt sich jedoch eine deutliche

17 Kommerziell erhältliche Anwendungen verfügen über eine kleinere Zahl an Elektroden und bestehen aus elektrolytgetränkten Filzstückchen, sodass sie mit geringem Aufwand zu montieren und zu unterhalten sind (dazu und zum Folgenden Hofmann 2014, S. 36; vgl. Hairston et al. 2014). Alternativ werden zunehmend sogenannte Trockenelektroden aus Golddrahtbürstchen oder auch kontaktlose, sogenannte kapazitive Elektroden eingesetzt. Beide Varianten sind in der Nutzung zwar einfacher, aber durch ihre unterschiedlichen Messmechanismen und wegen »ungewohnter« Fehlerquellen noch nicht als Ersatz für die Nasselektroden akzeptiert.

Verschlechterung der räumlichen Auflösung gegenüber epiduralen Ableitungen. Eine einzelne EEG-Elektrode erfasst die synchrone Aktivität von ca. 70 Mio. Zellen, während eine übliche ECoG-Elektrode immerhin ca. 1 Mio. Zellen abbildet.¹⁸ EEG-Signale sind zudem störanfällig und merklich verzerrt durch Schädelknochen und Kopfhaut, sodass ihre Interpretation selten eindeutig ist (dazu und zum Folgenden Hofmann 2014, S.22 u. 34 ff.). Trotzdem dient das EEG seit Dekaden als sehr erfolgreiches Diagnosehilfsmittel in der neurologischen Klinik (etwa zur Feststellung des Hirntods oder der Epilepsie). Dabei macht man sich zunutze, dass sich bestimmte Wellenmuster in EEG-Signalen offenbar grob lokalen Hirnfunktionen zuordnen lassen. Bei der Analyse eines EEG werden die Wellen daher zuerst in verschiedene Frequenzbänder zerlegt, deren Verteilung anschließend interpretiert wird (Kasten). Besonders interessant im Hinblick auf neurotechnologische Anwendungen ist, dass es offenbar auch Wellen gibt, die von Probanden mehr oder weniger erfolgreich gezielt erzeugt werden können – dieser Umstand lässt sich nicht nur therapeutisch nutzbar machen (sogenanntes Neurofeedback), sondern eröffnet auch die prinzipielle Möglichkeit zur Steuerung einfacher externer Maschinen mittels Hirnströmen (Kap. III.1.2).

Anders als beim Gehirn (mittels EEG) und bei Muskeln (mittels Elektromyografie, EMG) lassen sich elektrische Aktivitätssignale aus dem peripheren Nervensystem nicht an der Körperoberfläche gewinnen (dazu und zum Folgenden Hofmann 2014, S.39). Verantwortlich dafür ist der komplexe Aufbau peripherer Nerven, die eine große Zahl von Nervenfasern auf- und absteigender Signalrichtung aus unterschiedlichen Körperregionen bündeln. Dadurch ergibt sich auf der Hautoberfläche kein sinnvoll interpretierbares Summensignal. Insofern ist man bei der Ableitung von Signalen aus dem peripheren Nervensystem – der sogenannten *Elektroneurografie* – auf invasive Methoden angewiesen. Meist greift man dazu auf feine elektrisch leitende Nadeln zurück, die in die zähe, schützende Nervenhaut, das Epineurium, eingestochen werden. Der komplexe Aufbau peripherer Nerven, die neben einer großen Zahl verschiedener Nervenfasern auch Blutgefäße enthalten, erschwert jedoch auch hier ein präzises Vorgehen.

Neben den erwähnten Ableitungsverfahren, welche die neuronale Aktivität auf direktem Wege – also über die Erfassung elektrischer Erregungsleitung – abbilden, gibt es eine Reihe von Bildgebungsverfahren, die die Gehirnfunktion auf indirektem Wege messen. Hierzu dient vor allem die Bestimmung des Blutflusses auf lokaler Ebene (bzw. Veränderungen hiervon), da die Aktivität des Gehirns wie jede andere körperliche Betätigung auch von der Zufuhr von Sauerstoff und anderen Nährstoffen abhängig ist.

¹⁸ Normale Elektrokortikogramme haben eine Auflösung von etwas über 1 Mio. Neuronen, bei einem Mikro-ECoG (Elektrodendurchmesser von 0,8 statt 4 mm) reduziert sich dieser Wert noch einmal deutlich auf 50.000 Neuronen (Hofmann 2014, S.45).

EEG-Frequenzbänder und ihre funktionale Interpretation

EEG-Signale sind das Resultat der synchronen Aktivität größerer kortikaler und subkortikaler Nervenverbände. Die Frequenzverteilung eines EEG zeigt, aus welchen grundlegenden Wellenzügen das Signal aufgebaut ist, was Rückschlüsse auf spezifische Hirnaktivitäten (im Bereich der Großhirnrinde) erlaubt. Historisch hat es sich eingebürgert, zwischen dem δ -Band (2–4 Hz), dem θ -Band (4–8 Hz), dem α -Band (etwa 8–13 Hz), dem β -Band (etwa 13–30 Hz) und dem γ -Band (über 30 Hz) zu unterscheiden. Aus langjährigen Forschungen ist bekannt, dass diese Bänder grob mit bestimmten psychischen Prozessen korrespondieren (Niedermeyer/da Silva 2005).

So weiß man z. B., dass eine hohe Amplitude im δ -Band vor allem während des Schlafes im EEG zu sehen ist. Findet sich dieses Band sehr prominent während des Tages, kann das auf einen Hirnschaden hindeuten. θ -Band-signale werden durch die Wechselwirkung zwischen dem Frontalen und dem Temporalen Hirnlappen erzeugt. Sie tauchen während Einschlaf- bzw. Aufwachphasen auf und werden mit geringer Aufmerksamkeit bei gleichzeitiger Erinnerungskonsolidierung oder bildorientierter kreativer Aktivität in Verbindung gebracht. Die α -Wellen um 10 Hz erscheinen im EEG am besten sichtbar bei geschlossenen Augen im entspannten Zustand über dem Okzipitallappen. Funktional wird dieses Band mit Aufmerksamkeit bei fehlender kognitiver Verarbeitung assoziiert. β -Rhythmen erscheinen bei Menschen, die sich stark konzentrieren, vor allem im Frontal- und Parietallappen. Über dem Motorkortex zeigen sich α - und β -Wellen, auch bei der mentalen Vorstellung, man würde Hände, Füße oder die Zunge bewegen (Pfurtscheller et al. 1997). Ebenfalls leicht kontrollierbar und damit nutzbar scheinen die sogenannten μ -Rhythmen (8–12 Hz) und Teile des β -Bandes (15–25 Hz) zu sein (Pfurtscheller/Neuper 2001). Die γ -Wellen (über 30 Hz) bringt man vor allem mit dem Lösen kognitiver Aufgaben und höchster Konzentration in Verbindung.

Quelle: Hofmann 2014, S. 37 f.

Zur Darstellung der Durchblutungsänderung wurden verschiedene Verfahren entwickelt, die hier der Vollständigkeit halber Erwähnung finden sollen, obwohl sie bei neurotechnologischen Anwendungen derzeit eher noch eine Randrolle spielen (zum Folgenden Hofmann 2014, S. 52 ff.):

- > Die Nahinfrarotspektroskopie (NIR) basiert auf dem Umstand, dass das Blut je nach Sauerstoffgehalt eine unterschiedliche Färbung aufweist und insbesondere auch nahinfrarotes Licht unterschiedlich absorbiert. Da Licht in diesem Spektralbereich Schädel und Hirngewebe problemlos durch-



dringt, lässt sich mittels Durchleuchtung des Schädels mit nahinfrarotem Licht und Bestimmung des reflektierten Lichts indirekt auf die Hirnaktivität schließen. Die Methode ist ohne größeren technologischen Aufwand anwendbar und relativ kostengünstig, tiefere Gehirnregionen jenseits der Großhirnrinde lassen sich damit aber nicht abbilden.

- › Bei der Positronenemissionstomografie (PET) werden lebenswichtige Nährstoffe mit kurzlebigen radioaktiven Isotopen markiert und dem Probanden injiziert. Über den Blutfluss verteilen sich die Isotope im Gehirn, vornehmlich dort, wo der größte Nährstoffbedarf und also in der Regel auch die größte Aktivität vorherrschen. Aus der zeitlichen und räumlichen Verteilung des radioaktiven Zerfalls ergibt sich so indirekt ein grobes Bild der Hirnaktivität. Die PET-Untersuchung ist technologisch relativ aufwendig und entsprechend teuer.
- › Als Goldstandard bei den funktionellen bildgebenden Verfahren gilt seit einigen Jahren die funktionelle Kernspintomografie (fMRT). Hierbei macht man sich die unterschiedlichen magnetischen Eigenschaften von oxygeniertem und desoxygeniertem Blut zunutze, die sich in einem starken Magnetfeld, wie es von modernen Kernspintomografen erzeugt wird, sichtbar machen lassen (sogenannter BOLD-Kontrast [Blood Oxygenation Level Dependent]). Über das BOLD-Signal lassen sich detaillierte Rückschlüsse auf die Hirnaktivität ziehen, und zwar auch von tiefliegenden Strukturen. Das fMRT-Verfahren erfreut sich entsprechender wissenschaftlicher Beliebtheit, trotz des großen technologischen Aufwandes und der hohen Kosten einer Untersuchung. Allerdings sind die illustrativen fMRT-Bilder, die sicherlich viel zur Hochkonjunktur der Neurowissenschaften der letzten Jahre beigetragen haben, nur scheinbar eindeutig zu interpretieren, sondern das Resultat komplexer statistischer Analysen und mit entsprechend vielen methodischen Fallstricken behaftet.

Stimulationsverfahren

Da Nervenzellen grundsätzlich bioelektrisch aktiv sind, kann Strom auch zur Stimulation eingesetzt werden. Werden Neuronen einem elektrischen Feld ausgesetzt, verhält sich das Zytoplasma im Inneren wie ein elektrischer Leiter und generiert einen Strom entlang des Axons, der zur Depolarisation der Zelle führt (dazu und zum Folgenden Hofmann 2014, S.23 ff.). Je nach Stärke des elektrischen Feldes kann auf diese Weise ein Aktionspotenzial ausgelöst – falls der Schwellenwert überschritten wird – oder zumindest durch unterschwellige Reizung eine erhöhte Erregbarkeit des Neurons erreicht werden (sogenannte Langzeitpotenzierung). Da auf diese Weise im Grunde jedes Körperorgan stimuliert werden kann, gibt es eine große Bandbreite unterschiedlicher Stimulationsver-



fahren, die sich nach ihrem Wirkort und dem Grad ihrer Invasivität differenzieren lassen – das Spektrum reicht von der nichtinvasiven Stimulation von Muskeln (funktionelle Elektrostimulation), über den invasiven Herzschrittmacher bis hin zu der hochgradig invasiven Stimulation von Rückenmark und tiefen Hirnarealen. Aus neurotechnologischer Sicht sind dabei vor allem solche Ansätze von Interesse, die direkt auf das Gehirn einwirken und damit quasi am Zentrum der Reiz- und Informationsverarbeitung ansetzen.

Die heilende Wirkung extern verabreichter Stromschläge auf das Gehirn war offenbar bereits in der Antike bekannt (dazu und zum Folgenden Hofmann 2014, S.26 f. u. 61). Der römische Arzt Scribonius Largus beschreibt in seinen überlieferten Schriftstücken die Nutzung des Atlantischen Zitterrochens zur Behandlung unerträglicher Kopfschmerzen. Der Fisch sollte so lange auf die betroffene Körperstelle aufgelegt werden, bis diese taub war oder der Schmerz aufgehört hatte. Das Prinzip, dass durch ein starkes, extrakorporal erzeugtes elektrisches Feld neuronale Veränderungen im Gehirn bewirkt werden können, machen sich auch moderne nichtinvasive Stimulationsverfahren zunutze. Ein Beispiel ist die klinisch sehr effektive *Elektrokonvulsionstherapie* (EKT). Dabei wird unter kontrollierten Laborbedingungen und nach strenger Indikationsstellung durch eine starke elektrische Stimulation des Gehirns ein gezielter epileptischer Anfall im seitlichen Temporallappen ausgelöst, um durch die vorhandenen Selbstheilungskräfte eine Verbesserung des Neurotransmitterhaushalts zu erreichen. Um die bei epileptischen Anfällen auftretenden Muskelkrämpfe zu vermeiden, wird der betroffene Patient für kurze Zeit anästhesiert, während mit einem Muskelrelaxanz den Krämpfen vorgebeugt wird. Diese grobschlüchtig anmutende Prozedur erzielt bei geeigneter Indikation, z. B. anderweitig nicht behandelbaren Depressionen, anerkannt positive, wenn auch nur vorübergehende Resultate.

Geringere, dafür aber direkt im Gehirn erzeugte Feldstärken lassen sich durch die *transkranielle Magnetstimulation* (TMS) erreichen (dazu und zum Folgenden Hofmann 2014, S.26 u. 66 ff.). Anstatt über Elektroden am Kopf elektrische Impulse von außen zu verabreichen, wie bei der EKT, wird hier durch Spulen ein Magnetfeld direkt im Schädelinneren aufgebaut. Durch die schnelle Änderung des Magnetfeldes induziert man in den oberen Zentimetern der Hirnrinde Wirbelströme, die wiederum die dortigen Neuronen aktivieren können. Abhängig davon, ob es sich um eine einmalige Stimulation handelt (sTMS) oder eine repetitive Serie von niederfrequenten Stimulationspulsen (rTMS), hat die Anwendung unterschiedliche Auswirkungen auf das stimulierte Netzwerk. So lassen sich mittels einer einmaligen Stimulation (sTMS) des Motorkortex Aktionspotenziale entlang der Nervenfasern und in der Folge periphere Muskelkontraktionen erzeugen (Hallett 2000), eine Methode, die zur Diagnose motorischer Störungen wie der Multiplen Sklerose (MS), der Amyotrophen



Lateralsklerose (ALS) oder von Schlaganfall genutzt werden kann (Groppa et al. 2012). Dagegen erreicht man mit rTMS längerfristige und unspezifischere Auswirkungen, die sich in Aktivitätsänderungen im stimulierten Bereich der Großhirnrinde zeigen können. Inwiefern sich derartige Effekte therapeutisch nutzbar machen lassen – beispielsweise gegen schwere Depressionen und schwere Migräne –, wird derzeit erforscht. Da die magnetischen Felder senkrecht und ohne Dämpfung durch den Schädel dringen und erst im Hirngewebe Wirbelströme induzieren, ist die Methode grundsätzlich frei von oberflächlichen Schmerzen, wie sie durch senkrecht die Haut durchdringende Ströme auftreten können. Der klinische Nutzen der Therapie ist aber umstritten und scheint stark von den Stimulationsparametern abhängig zu sein (Herwig et al. 2007; O'Reardon et al. 2007). Aufgrund der Uneinheitlichkeit der Forschungsergebnisse und unklarer Nebenwirkungen gibt es eine klinisch-therapeutische und damit finanziell kompensierte Anwendung für TMS erst in Großbritannien und den USA, jedoch nicht in Deutschland. Die Komplexität der Technologie mit Hochspannungskondensatoren und gekühlten Hochstromspulen hat bisher auch dazu geführt, dass der Markt der transkraniellen Magnetstimulation von kommerziellen Anbietern versorgt wird und Heimentwicklungen noch keine große Rolle spielen.

Anders sieht dies bei der *transkraniellen Gleich- oder Wechselstromstimulation* (tDCS oder tACS) aus. Von den USA ausgehend, appliziert sich eine wachsende Do-it-yourself-Community schwache Ströme über die Kopfhaut im Selbstversuch – etwa, indem in Salzlösung getränkte Schwämme am Kopf befestigt und mit einer Batterie verbunden werden (dazu und zum Folgenden Hofmann 2014, S.63 ff.). Obwohl auf der einen Seite der Schädelknochen jedem Stromfluss durch seinen geringen Wassergehalt einen hohen Widerstand entgegengesetzt und auf der anderen Seite alleine schon durch die Sensibilität der Haut nicht beliebig hohe Stimulationsleistungen möglich sind, konnte in Studien gezeigt werden, dass derart verabreichte schwache Ströme im Bereich von 1 bis 2 mA durchaus eine unspezifische Wirkung auf das Gehirn ausüben (Nitsche/Paulus 2000). Es wurden unterschiedlichste Effekte beschrieben: Verbesserungen von Lern- und Merkfähigkeit, des Erinnerungsvermögens, der Behandlung von Migräne und Depression und sogar eine verringerte Kalorienaufnahme (Antal et al. 2011; Brunoni et al. 2012; Clark/Parasuraman 2014; Kaminski et al. 2013; Meinzer et al. 2014). Bereits halten vereinzelte Experten das Enhancement der Hirnfunktionen durch schwache Ströme für eine aussichtsreiche Zukunftsoption (Coffman et al. 2014), was etwas vorschnell erscheint, da es aktuell noch keine detaillierten und langfristigen Untersuchungen zu den tDCS/tACS-Auswirkungen auf das Gehirn gibt. Ein Bericht des Nuffield Council on Bioethics kam in diesem Zusammenhang im Jahr 2013 zu der Einschätzung, dass man sehr vorsichtig damit sein sollte, von den Ergebnissen einiger weniger kleiner Studien unter Laborbedingungen auf die Möglichkeit dauerhafter Effekte in der



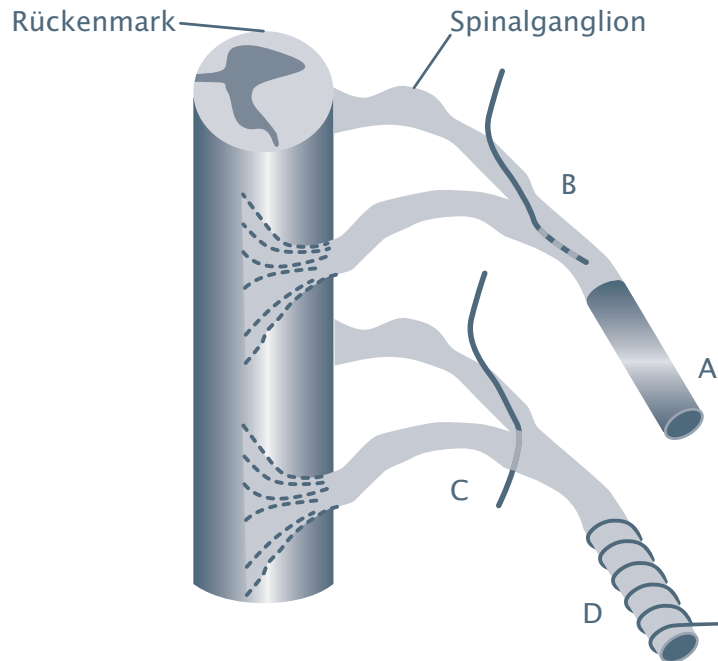
»realen Welt« zu schließen (Baldwin et al. 2013, S.165 ff.). Auf jeden Fall sind die in der Literatur berichteten, durchgängig eher geringfügigen Wirkungen nicht geeignet, weitreichende Visionen einer Verbesserung der menschlichen Natur zu rechtfertigen. Auch der eigentliche Wirkmechanismus ist noch unbekannt und Anlass für weitere Forschungsarbeit.

Wesentlich spezifischere Wirkungen als durch die beschriebenen unblutigen Verfahren können durch *invasive Stimulatoren* erreicht werden (dazu und zum Folgenden Hofmann 2014, S.69 ff.). Hierbei werden Elektroden direkt in die Nähe des zu stimulierenden Bereichs implantiert – sei es tief im Gehirn, in Muskeln, Rückenmark oder an peripheren Nerven –, sodass die Wirkung der Stimulation auf ein enger umgrenztes Gebiet beschränkt bleibt. Da in all diesen Fällen die körperliche Unversehrtheit des Empfängers verletzt wird, sind darauf basierende Anwendungen (etwa die tiefe Hirnstimulation, Kap. III.1.2) derzeit nur bei einer entsprechenden medizinischen Indikation durchführbar. Die Miniaturisierung der Stimulationssysteme durch geeignete Mikrosystemtechnik ist dabei ein zentrales Forschungsziel, denn die elektrische Schädigung des umliegenden Gewebes scheint bei mikroskopischen Stimulationselektroden offenbar geringer zu sein als bei Makroelektroden. Vor allem aber hat die Mikrostimulation den Vorteil, eine sehr selektive und damit nebenwirkungsarme elektrische Reizung zu ermöglichen. Die Selektivität ist vor allem bei der Stimulation peripherer Nerven, die aufgrund ihrer gebündelten Struktur schwer gezielt zu reizen sind, ein zentraler Aspekt. Lange behalf man sich mit einhüllenden, manschettenartigen Elektroden (sogenannte Cuff-Elektroden), die nur über wenige Kontakte verfügen (Abb. III.3). Erst seit jüngster Zeit lässt sich durch mikrotechnische Vielfachelektroden wie das Utah-Array eine größere Anzahl von Kontakten herstellen. Trotzdem bringen auch die derzeit am höchsten entwickelten Sonden zur Stimulation peripherer Nerven nur eine Handvoll Elektroden in oder um die jeweilige große Nervenfasern ein, sodass noch keine ausgeprägt hohe Auflösung und Spezifität zu erwarten ist.

Ein grundsätzliches Problem implantierter Schnittstellen (sowohl stimulierender als auch ableitender) ist die *Fremdkörperreaktion* (zum Folgenden Hofmann 2014, S.29 ff. u. 85). Ihr liegt eine Immunantwort des Körpers zugrunde, die zur Einkapselung der Implantate mit Narbengewebe führt, sodass neuroelektrische Schnittstellen (insbesondere auf mikroskopischer Skala) den Kontakt zum neuronalen Gewebe verlieren und damit nach kurzer Zeit ihre Funktion einbüßen. Die Sicherung der langfristigen Stabilität und Funktionalität implantierter Mikrosysteme über Jahre und Jahrzehnte ist deshalb ein wichtiges Ziel und gleichzeitig eine der größten Herausforderungen für die Forschung, die man durch die Entwicklung neuer biokompatibler Materialien und Oberflächenstrukturen angeht (Nguyen et al. 2014). Der vielversprechendste Ansatz weist im Moment weg von rigiden Implantaten wie Metalldrähten oder Silizi-

umnägeln hin zu flexiblen und weichen Implantaten, die im Gewebe mit-schwimmen und weniger eingekapselt werden. Möglichem oxidativen Stress durch das fremde Material versucht man u. a. durch Beschichtung mit Materialien, die selbst sehr viel Wasser enthalten (sogenannte Hydrogele), zu begegnen.

Abb. III.3 Visualisierung elektrischer Kontaktierung am peripheren Spinalnerv



A) Cuff-Elektrode, B) LIFE-Elektrode, C) TIME-Elektrode, D) Spiral- oder Helixelektrode

Quelle: nach Hofmann 2014, S. 72

In den letzten 10 Jahren wurde mit der sogenannten *Optogenetik* ein ganz neuartiges Stimulationsverfahren entwickelt, das die Nervenzellen auf optischem (statt elektrischem) Wege stimuliert (dazu und zum Folgenden Hofmann 2014, S.80 ff.). Dieses neue Verfahren wurde in Fachkreisen euphorisch aufgenommen und gilt als »Quantensprung« für die neurowissenschaftliche Forschung, weil auf diese Weise eine ganz gezielte Lichtstimulation spezifischer Neuronentypen in bestimmten Hirnregionen möglich ist. Dies eröffnet bislang ungeahnte medizinische Perspektiven, etwa die Stimulierung elektrisch nicht erreichbarer Hirnregionen zu therapeutischen Zwecken. Bis zur Realisierung ist es jedoch noch ein weiter Weg, denn das Verfahren setzt die gentechnische Manipulation der zu beeinflussenden Nervenzellen voraus, die mit lichtempfindlichen Ionenkanälen ausgestattet werden müssen. Aus diesem Grunde ist das Verfahren bislang nur von experimentellem Wert und hat beim Menschen noch auf längere Sicht keine konkreten Anwendungschancen.

Neurotechnologische Anwendungen: Stand und Entwicklungsperspektiven

1.2

Auf Basis der zuvor beschriebenen neuroelektrischen Schnittstellen werden seit vielen Jahren neurotechnologische Anwendungen konzipiert und auch entwickelt, die technische Systeme an Nerven des peripheren oder zentralen Nervensystems koppeln. Je nachdem, in welche Richtung die Signalübertragung stattfindet, sind diese Anwendungen ableitenden oder stimulierenden Typs – abhängig davon, ob Hirn- bzw. Nervensignale dazu benutzt werden, externe Apparaturen zu steuern (z. B. einen Computer), oder ob umgekehrt Hirn- bzw. Nervenaktivität durch elektrische Reize technisch beeinflusst wird. Zunehmend geraten auch bidirektionale Anwendungen in den Fokus von Forschung und Entwicklung. Hierbei wird versucht, durch Integration von ableitenden und stimulierenden Schnittstellen ein geschlossenes Feedbacksystem zu etablieren, beispielsweise im Rahmen fühlender Greifprothesen, die Sensorik und Aktorik verbinden. So ergibt sich eine große Bandbreite an möglichen neurotechnologischen Anwendungen, die sich in unterschiedlichen Phasen der Entwicklung bzw. Anwendungsreife befinden (Tab. III.1).

Tab. III.1 Kategorisierung neurotechnologischer Anwendungen

Richtung der Signalübertragung	Art der Schnittstelle	
	nichtinvasiv	invasiv
ableitend	Brain-Computer-Interfaces (z. B. Thought Translation Device) Targeted Muscle Reinnervation (TMR)	Brain-Machine-Interfaces (z. B. komplexe Greifprothesen)
stimulierend	Fallfußstimulator	tiefe Hirnstimulation sensorische Neuroprothesen (Cochlea- und Retina-Implantat etc.) Rückenmarkstimulation
bidirektional		Elektrozeptika fühlende Greifprothese

Eigene Zusammenstellung

Ableitende Systeme

Trotz sehr unterschiedlicher Spielarten basieren alle ableitenden neurotechnologischen Anwendungen auf dem gleichen Grundprinzip. Die neuronale Aktivität, welche die Intention des Nutzers kodiert, wird im peripheren bzw. zentralen Nervensystem durch geeignete elektrophysiologische Schnittstellen¹⁹ erfasst, digitalisiert und an einen Computer weitergeleitet. Dort werden die Signale analysiert, um die Intention des Nutzers zu dechiffrieren, in Steuerbefehle übersetzt und schließlich über eine Programmschnittstelle an die jeweilige Anwendung übermittelt. Meist findet dabei noch eine sensorische Rückkopplung statt, zumeist in visueller Form, durch die der Nutzer lernt, das System besser zu kontrollieren (Abb. III.4).

Zur Bewältigung dieser zahlreichen Prozessierungsschritte ist ein komplexer apparativer Aufbau erforderlich, der verschiedene Hardware- und Softwarekomponenten umfasst, die sich im bzw. am Körper (je nach Art der Schnittstelle), zum Großteil jedoch außerhalb des Körpers befinden. Neben der Qualität und Verträglichkeit der ableitenden Sensoren hängt der Erfolg eines solchen Systems vor allem auch davon ab, inwiefern es gelingt, aus den riesigen erfassten Datenmengen mittels geeigneter Algorithmen die relevanten, teilweise hochindividuellen Gehirnmuster zu erkennen und zu klassifizieren. Für diesen Prozess werden zunehmend Algorithmen eingesetzt, die auf Methoden des maschinellen Lernens basieren.

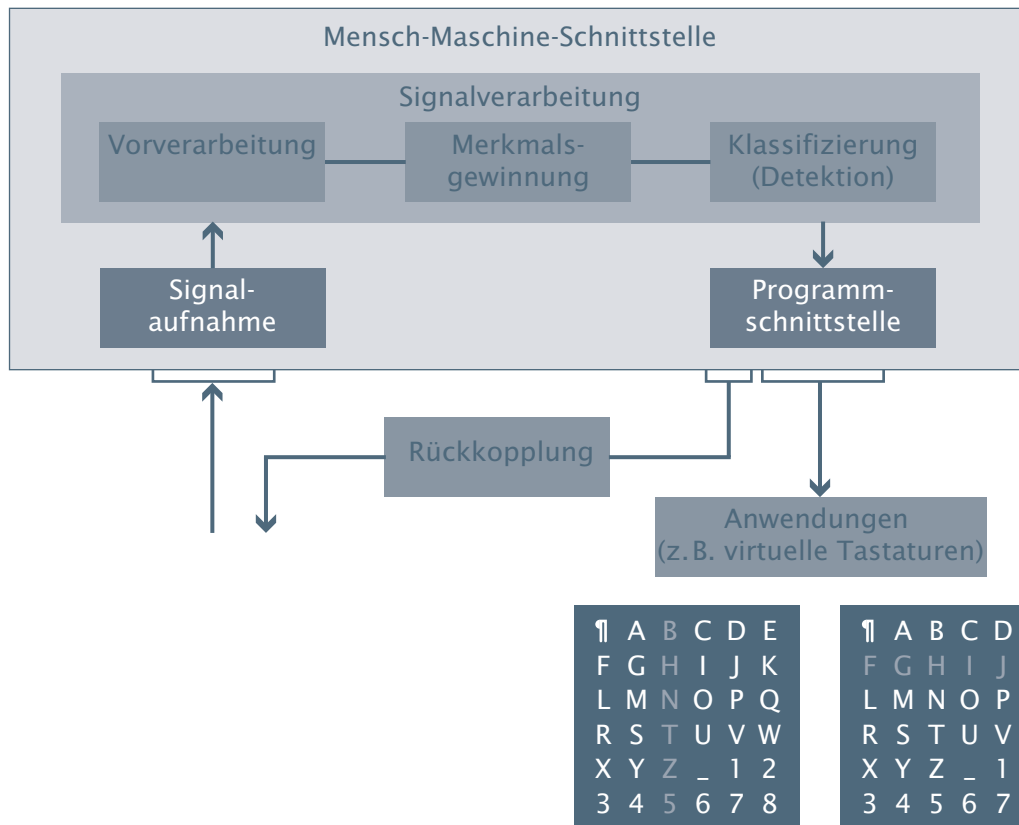
Ableitende Systeme unterscheiden sich danach, wo (im peripheren oder zentralen Nervensystem) und auf welche Art (invasiv vs. nichtinvasiv) neuronale Aktivität erfasst wird, sowie natürlich auch danach, welche Geräte letztendlich gesteuert werden (Prothese, Rollstuhl, Computercursor etc.).²⁰ Die Prothesensteuerung ist einer der zentralen und angesichts der Vielzahl der Betroffenen auch kommerziell aussichtsreichsten Anwendungsbereiche. Seit über 30 Jahren behilft man sich hier mit myoelektrischen Prothesen, bei denen Elektroden die verbliebene Muskelaktivität im Stumpf erfassen und damit batteriebetriebene Motoren in Gang setzen. Besonders bei der technischen Nachbildung der Hand mit ihren verschiedenen Bewegungsmöglichkeiten (Öffnen/Schließen, Drehen,

19 Alternative Herangehensweisen wie bildgebende Verfahren (fMRT etc.) spielen derzeit eine untergeordnete Rolle und werden dies vermutlich auch noch auf längere Zeit tun, da sie einen weit höheren maschinellen Aufwand erfordern.

20 Wolpaw (2014) unterscheidet in diesem Zusammenhang fünf Anwendungszwecke von Brain-Computer-Interfaces: das Ersetzen einer Funktion (replace; z. B. die Steuerung eines Rollstuhl), die Wiederherstellung einer Funktion (restore; z. B. die Steuerung einer Prothese), das Bereitstellen steigernder (enhance; z. B. die Überwachung der Aufmerksamkeit), ergänzender (Supplement; z. B. ein zusätzlicher Roboterarm) oder verbessernder (improve; z. B. die Unterstützung der Schlaganfalltherapie) Funktionalitäten.

Abb. III.4

Grundaufbau ableitender Anwendungen



Quelle: nach Stieglitz et al. 2014, S. 459

Abwinkeln) stößt diese Methode aber schnell an Grenzen. Aufgrund der geringen Spezifität der Daten erlauben herkömmliche myoelektrische Prothesen in der Regel nur eine sequentielle Steuerung der unterschiedlichen Bewegungsfunktionen und sind deshalb nur (sehr) eingeschränkt praktikabel (Graumann 2015, S. 21 f.). Es gibt in diesem Bereich in den letzten Jahren aber einige Fortschritte zu verzeichnen. Zum einen wird an »intelligenten« Roboterprothesen geforscht, die mit Sensoren u. a. zur Umwelt- und Objekterkennung ausgestattet sind und dank ausgefeilter Software eine teilautonome Unterstützung des Greifprozesses ermöglichen sollen (Graumann 2015, S. 24). Dies könnte die Handhabung komplexer Greifprothesen für den Nutzer deutlich erleichtern. Bereits in der klinischen Praxis etabliert ist zum anderen ein 2002 entwickeltes Verfahren, das als *selektiver Nerventransfer* (Targeted Muscle Reinnervation [TMR]) bekannt ist. Es ermöglicht armamputierten Patienten die Steuerung einer Prothese über den Brustmuskel, indem dieser quasi als Verstärker der schwachen nervösen Signale dient, die vorher der Bewegung des Armes dienten



(dazu und zum Folgenden Hofmann 2014, S. 40 ff.). Dazu werden die verbliebenen Armnerven in verschiedene Segmente des zuvor deinnervierten Brustmuskels umgeleitet, wo sie wieder aussprossen und Informationen übertragen. Die so entstehenden muskulären Potenziale können dann mittels Elektrodenableitung erfasst und zur Steuerung eines kompletten roboterisierten Armes mit Greifhand benutzt werden. Armprothesen, die sich dieses Prinzip zunutze machen, sind bereits auf dem Markt.²¹ Obwohl die eigentlich nichtinvasive Ableitung des reinnervierten Brustmuskels erst durch eine invasive und risikobehaftete Operation möglich wurde, ist zu erwarten, dass diese Technik im Laufe der kommenden zwei Dekaden zum Goldstandard bei Vollamputationen der Extremitäten aufsteigt.

Entscheidend für den Erfolg dieses Systems ist, dass Patienten keine elaboreierten Kontrollsignale lernen müssen, sondern die Prothese nach geringem Trainingsaufwand intuitiv durch gedankliches Armbewegen steuern können (Hofmann 2014, S. 42; Kuiken et al. 2009). Das gelingt allerdings nur dann, wenn die Hirnsignale noch das periphere Nervensystem erreichen, was bei komplett Gelähmten nicht mehr der Fall ist. Außerdem lassen sich bei TMR durch das gezielte Umleiten der Armnerven zwar durchaus natürliche Bewegungsabläufe erreichen – nach aktuellem Stand der Technik sind drei Gelenke gleichzeitig ansteuerbar, sodass sich z. B. mit etwas Übung ein Ball aus der Luft fangen lässt (Kuiken et al. 2009) –, allerdings wird das Verfahren derzeit nur bei Armamputation im Bereich des Schultergelenks eingesetzt.²² Aus diesen Gründen erscheinen ableitende Systeme, welche die Signale direkt aus dem Gehirn gewinnen, oft erfolgversprechender, wenn es darum geht, Prothesen oder andere komplexe Gerätschaften über Gedankenkraft zu steuern. Die ersten Gehirn-Maschine-Schnittstellen, die in den 1990er Jahren entwickelt worden waren, basierten alle auf EEG-Ableitungen, waren also nichtinvasiv, und dienten vornehmlich als computerbasierte Kommunikationshilfe. Bei den später aufkommenden invasiven Varianten steht hingegen die Steuerung komplexer Prothesen im Vordergrund. Aus diesem Grund werden – wie es auch in Teilen der Fachliteratur üblich ist – im Kontext dieses Berichts Anwendungen der ableitenden Art, die nichtinvasive Schnittstellen beinhalten, als *Brain-Computer-Interfaces* (BCI) bezeichnet, solche mit invasiven Schnittstellen als *Brain-Machine-Interfaces* (BMI).

21 <https://professionals.ottobock.ca/zb2b4ob/us01/en/USD/Prosthetics/Upper-Limb-Prosthetics/Myoelectric-Elbows/DynamicArm-TMR/p/12K110~550> (15.12.2015)

22 Dies hat damit zu tun, dass im Prothesenschaft die hochempfindlichen Elektroden angebracht sind und die Prothese für eine optimale Signalqualität direkt am Oberkörper befestigt wird. Die Eignung des TMR-Verfahrens zur Steuerung von Beinprothesen wird derzeit erforscht (Stieglitz 2015, S. 12).

Brain-Computer-Interfaces

Bis heute setzt der Großteil der BCI-Anwendungen auf EEG-Ableitungen als Methode der Wahl, um die Gehirnaktivität zu ermitteln. Nur vereinzelt werden dazu bildgebende Verfahren wie die funktionelle Magnetresonanztomografie (fMRT) verwendet, die sich aufgrund des hohen apparativen und finanziellen Aufwandes auch auf absehbare Zeit für eine Nutzung außerhalb des Labors als nicht praktikabel erweisen dürften. Das EEG hingegen hat den Vorteil, relativ einfach anwendbar zu sein und Kommunikation durch eine hohe zeitliche Auflösung zu ermöglichen (Kübler/Neuper 2008, S.208). Dabei macht man sich zunutze, dass ableitbare Gehirnpotenziale in bestimmten Frequenzbändern mit bestimmten psychischen Prozessen korrelieren, die vom Probanden bewusst steuerbar sind und somit zur Kontrolle eines externen Systems genutzt werden können (dazu und zum Folgenden Hofmann 2014, S.93 ff.):

- > *Steady-state visuell evozierte Potenziale*: Dabei nutzt man die Eigenschaft des visuellen Systems aus, dass es ein elektrisches Abbild der Außenwelt erstellt (Kelly et al. 2005). Lässt man unterschiedliche Bereiche eines Computermonitors mit unterschiedlichen Frequenzen blinken, wird man diese Blinkfrequenz im EEG-Signal wiederfinden. Aus der gemessenen Frequenz lässt sich folglich auf den Bereich des Schirms schließen, auf den der Proband geblickt hat. Verschiedene Buchstabierhilfen basieren auf diesem Phänomen.
- > *Ereigniskorrelierte Potenziale* werden durch die Präsentation visueller, taktiler oder auditorischer Stimuli ausgelöst. Das Signal-Rausch-Verhältnis dieser Kurven ist schlecht, sodass sich deren Identifikation im EEG als große Herausforderung darstellt, die meist durch wiederholte Präsentation des Reizes gelöst wird. Trotzdem basiert eine ganze Klasse von BCI, die zu Kommunikationszwecken genutzt werden, auf diesem Signal. Weit verbreitet sind die sogenannten P300-Speller, Buchstabiersysteme, die sich die gut nachweisbare P300-Welle zunutze machen. Diese zeigt sich etwa 300 ms, nachdem ein auslösender Reiz – in diesem Fall ein bestimmtes, zufällig eingeblendetes Buchstabensymbol – vom Nutzer wahrgenommen wurde.
- > *Langsame kortikale Potenziale* sind eine Spezialform ereigniskorrelierter Potenziale, die sich als »langanhaltende Verschiebungen des gesamten EEG-Spektrums« in Reaktion auf äußere, aber auch innere Reize äußern (Schneider/Strauß 2013). Nutzer können in langwierigem mentalem Training lernen, diese Potenziale mittels gedanklicher Vorstellungen gezielt in positiver oder negativer Richtung zu beeinflussen (Birbaumer 2006). Auf diesem Prinzip basierte eines der ersten BCI-Systeme, das »Thought Translation Device«, das in den 1990er Jahren von einer Gruppe um Nils Birbaumer an der Universität Tübingen entwickelt wurde (Birbaumer 2000). Schwerstge-



lähmten Patienten ist es damit gelungen, einen Computercursor zu steuern und Texte zu verfassen.²³

- › Auch das *Fehlerpotenzial* (Error-related Potential) ist im Grunde ein ereigniskorreliertes Potenzial. Es zeigt sich, sobald ein Nutzer bemerkt, dass das erzielte Resultat (z. B. die Wahl eines Buchstabens) nicht mit seinen Erwartungen übereinstimmt. Damit ist ein starkes weiteres Muster erkennbar, das sogenannte Fehlerpotenzial, mit dem die Eingabe korrigiert oder eine andere Aktion ausgeführt werden kann.

Wie diese Auflistung bereits andeutet, sind EEG-basierte BCI-Systeme bislang vor allem als computerbasiertes Kommunikationsmittel (Buchstabierhilfe bzw. Surfen im Internet) für schwer gelähmte Patienten intensiv und mit einigem Erfolg erprobt worden. Andere Anwendungszwecke – insbesondere die Steuerung externer Geräte wie Rollstühle, Hilfsroboter, Exoskelette, Drohnen etc. (Clausen 2010b, S. 9 ff.) – werden zwar seit vielen Jahren intensiv erforscht, bis zum Praxiseinsatz außerhalb des Labors unter Alltagsbedingungen ist es jedoch noch ein weiter Weg (Lotte et al. 2015).

Ein wichtiger Spezialfall stellt die Nutzung der BCI-Technologie zu therapeutischen Zwecken dar, das sogenannte Neurofeedback (Huster et al. 2014). Hierbei werden die per EEG aufgezeichneten Hirnsignale in eine Computeraanimation übersetzt (z. B. einfaches Steuern eines Flugzeugs), was es dem Patienten ermöglicht, seine Hirnaktivität nachhaltig in eine bestimmte Richtung zu beeinflussen. Ein derartiges Gehirntraining gilt als erfolgversprechende Alternative zu medikamentösen Therapien bei der Behandlung diverser psychischer Störungen (darunter insbesondere der Aufmerksamkeitsdefizit- bzw. Hyperaktivitätsstörung), inwiefern es auch bei der unterstützenden Rehabilitation von Hirnschädigungen weiterhilft (z. B. nach einem Schlaganfall), wird untersucht (Hammond 2007).

Das Feld der BCI-Systeme erweist sich folglich als florierender Forschungsbereich, entsprechende Anwendungen stecken aber, bis auf wenige Ausnahmen, noch in den Kinderschuhen (dazu und zum Folgenden Hofmann 2014, S. 96 ff.). Schnelle Erfolge sind auch in nächster Zeit nicht zu erwarten, was mit verschiedenen Faktoren zusammenhängt. Erstens ist die Nichtinvasivität zugleich größ-

23 Allerdings setzt die Kommunikation mittels EEG-basierter BCI-Systeme interessanterweise voraus, dass die Patienten noch über eine zumindest minimale Bewegungskontrolle verfügen, und sei es nur der Augen- oder Gesichtsmuskeln (Birbaumer/Chaudhary 2015). Bei vollständig Gelähmten helfen sie jedoch nicht weiter, obwohl gerade diese Patientengruppe am allermeisten auf diesen Kommunikationsweg angewiesen wäre. Erste Erfolge konnten jedoch mit einem neu entwickelten BCI-System erzielt werden, das nicht an der neuroelektrischen Hirnaktivität ansetzt, sondern den Blutfluss im Gehirn (bzw. dessen Veränderung) mittels Nahinfrarotspektroskopie bestimmt (Gallegos-Ayala et al. 2014).

ter Pluspunkt wie auch wesentlicher Nachteil der Technologie. Denn die Gewinnung von Oberflächenpotenzialen auf der Kopfhaut von gesunden wie kranken Nutzern ist zwar unproblematisch und risikolos, stellt gleichzeitig aber eine immense Herausforderung an die Signalverarbeitung und sinnvolle Nutzung dieser Daten dar. Aufgrund der Instabilität der Signale, ihrer Verrauschung mit Signalanteilen, die nicht aus dem Gehirn stammen,²⁴ und einer bemerkenswerten Variabilität zwischen Versuchsdurchgängen und zwischen Probanden gelten BCI-Applikationen noch als äußerst trainingsintensiv und wenig verlässlich in Nutzung und Leistung – was insbesondere außerhalb des Labors und in der Daueranwendung ins Gewicht fällt. Zweitens bieten die zuvor aufgeführten Verfahren in der Regel nur eine binäre Antwort – entweder zeigt sich ein bestimmtes Potenzial oder nicht –, und sind damit schwerlich geeignet, komplexere Anwendungen zu bedienen, die kontinuierliche Kontrollsignale benötigen (z. B. für die fein abgestimmte Kontrolle von Geschwindigkeit oder Kraftausübung). Drittens stellt schließlich die geringe Bandbreite der übertragbaren Steuerkommandos einen Engpass dar. Eine Übertragungsrate von ein bis zwei Zeichen pro Minute (Schreuder et al. 2011) erscheint für die Steuerung komplexer, zeitkritischer Geräte als sehr gering – zumindest für den realen Einsatz –, und sie dürfte sich auf absehbare Zeit auch kaum wesentlich steigern lassen.²⁵ Klar ist also, dass die BCI-Technologie noch großer Entwicklungssprünge bedarf, um das enge Feld der laborgetragenen BCI zu verlassen und eine weite Verbreitung zu erreichen. Neben der Verbesserung der Leistungsfähigkeit der Elektrodenschnittstellen richten sich die Entwicklungsbemühungen derzeit vor allem auf sogenannte hybride BCI-Systeme (Fazli et al. 2012), die verschiedene BCI-Signale oder auch andere Signalquellen (z. B. Muskelsignale) nutzen, um dadurch mehrere unabhängige Steuerkanäle zu erhalten. Außerdem greift die BCI-Forschung zunehmend Erkenntnisse und Ergebnisse der autonomen Robotik auf (Lampe et al. 2014). Das Ziel ist, die sparsamen Steuersignale einem maximalen Nutzen zuzuführen, indem die »Intelligenz« autonomer Systeme gezielt eingesetzt wird (del R. Millán et al. 2010). Diese Forschungen befinden sich jedoch noch in einem frühen Stadium.

Obwohl sich die Entwicklung von BCI-Anwendungen derzeit noch fest in der Hand professioneller Neurowissenschaftler und klinischer Psychologen befindet, sind erste Bestrebungen zu beobachten, die Technologie einer breiteren Nutzergruppe verfügbar zu machen (dazu und zum Folgenden Hofmann 2014, S. 98 f.). Die zuvor erwähnte Neurofeedbacktherapie verbreitet sich zunehmend in psycho- und ergotherapeutischen Praxen. Verschiedene Hersteller bieten da-

24 Problematisch sind hier insbesondere Störeinflüsse, die durch die Bewegung von Gesichtsmuskeln entstehen, weshalb die Probanden in einem wissenschaftlichen Experiment in der Regel gebeten werden, Kopf und Gesicht möglichst stillzuhalten.

25 Der derzeitige Rekord steht bei etwa zehn Zeichen pro Minute (Bin et al. 2009).



zu auf kommerziellem Wege vereinfachte, relativ preiswerte und damit massentaugliche EEG-Systeme an (so etwa die Firmen emotiv und Brainfinger) (Kaplan et al. 2013; Ahn et al. 2014), die auch eine langsame, aber handfreie Navigation durch Computerspiele ermöglichen sollen (Hazrati/Hofmann 2013). Gleichzeitig ist mit BCI2000 eine mächtige und kostenlose Softwaretoolbox verfügbar, mit der auch Anfänger in die Lage versetzt werden, BCI-Experimente durchzuführen (Schalk et al. 2004).²⁶ Somit ist es wohl nur eine Frage der Zeit, bis auch dieses Feld von Do-it-yourself-Anwendern erobert wird, die sicherlich ganz neue Anwendungsmöglichkeiten jenseits des klinisch-therapeutischen Spektrums erkunden werden. Allerdings ist nicht nur die Zuverlässigkeit kommerziell erhältlicher Systeme noch sehr begrenzt, sondern es stellt sich auch die prinzipielle Frage, welcher Zusatznutzen von BCI-Systemen überhaupt für gesunde Personen zu erwarten ist, die sich im Vollbegriff ihrer körperlichen Kräfte befinden. Denn bestenfalls gewinnt der Nutzer die Leistungsfähigkeit einer klemmenden Maustaste hinzu, sodass die Steuerung von Geräten per Muskelkraft praktisch in allen Anwendungskontexten voraussichtlich noch auf längere Zeit bei Weitem effektiver, schneller und zuverlässiger sein wird.

Brain-Machine-Interfaces

Auf der Suche nach schärferen und höher aufgelösten Signalen sind einige BCI-Forscher dazu übergegangen, ihre Datenquellen um epidurale Kortexableitungen (ECoG) zu erweitern. Dank der höheren und spezifischeren Datenraten, die sich auf diese Weise gewinnen lassen, eröffnen sich ganz neue Möglichkeiten der Gerätesteuerung (Ortiz-Catalan et al. 2012) – jedoch geht dies nur auf Kosten einer hochriskanten invasiven Prozedur, womit das Grunddilemma der per definitionem »blutigen« BMI-Systeme (zu denen auch ECoG-Systeme gehören) bereits umrissen ist.

Das große Ziel, das mit invasiven im Gegensatz zu nichtinvasiven Schnittstellen durchaus erreichbar scheint, ist die Entwicklung komplexer robotischer Gliedmaßen, die sich auch von schwerstgelähmten Patienten intuitiv steuern lassen. Intuitiv heißt, dass die Steuerung nicht indirekt erfolgt, über die bewusste Beeinflussung bestimmter Hirnfrequenzen (wie im Falle von BCI-Anwendungen), sondern direkt über Gedankenkraft. Hierzu kommen jedoch Oberflächenableitungen nicht infrage, die auch im Falle von ECoG immer noch große Neuronenverbände abgreifen, sondern es braucht vielmehr zielgenaue Ableitungen entweder aus dem peripheren Nervensystem oder aus dem Motorkortex. Weil es große Schwierigkeiten macht, auf invasivem Wege eine stabile und leistungsfähige Schnittstelle zu den äußerst feinen peripheren Nerven aufzubauen,

²⁶ www.schalklab.org/research/bci2000 (15.12.2015)



scheint der direkte Anschluss ans Gehirn derzeit der gangbarere Weg zu sein. Dank moderner Mikroelektroden wie dem Utah-Array lassen sich dort Einzelzellaktivitäten von mehreren Zellen simultan erfassen, womit sich – wie in verschiedenen Tierversuchen gezeigt werden konnte – Bewegungsabsichten recht akkurat vorhersagen lassen (dazu und zum Folgenden Hofmann 2014, S. 100 f.). In einer wachsenden Zahl von Affenexperimenten ist es so seit Mitte der 2000er Jahre gelungen, die prinzipielle Machbarkeit eines BMI für prothetische Zwecke zu belegen (Carmena et al. 2003; Chapin et al. 1999; Nicolelis/Chapin 2008; Schwartz 2004; Wessberg et al. 2000). Versuchstiere lernten, verschiedene roboterisierte Manipulatoren wie Gliedmaßen akkurat zu kontrollieren, wenngleich die Steuerung aufgrund fehlender sensorischer Rückmeldung nur über Sichtkontakt funktioniert. Einige Gruppen berichten aktuell von erfolgreichen BMI-Ableitungen aus Makaken mit vollimplantierten, kabellosen BMI-Systemen, die über viele Monate verlässlich einen Datendurchsatz von 4bit/sec liefern (Ches-tek et al. 2011; Churchland et al. 2012; Nuyujukian et al. 2014).

Da ein komplexer operativer Eingriff erforderlich ist, um die Sonden zu implantieren, und vor allem die durch die Schädeldecke führende Verkabelung hohe Infektionsrisiken birgt, wurden prothetische BMI bislang nur in Einzelfällen am Menschen erprobt (dazu und zum Folgenden Hofmann 2014, S. 101 ff.). 2006 berichtete erstmals eine Gruppe der US-amerikanischen Brown University von der erfolgreichen Implantation und Nutzung eines Utah-Arrays im Motor-kortex von Tetraplegikern (Hochberg et al. 2006). In jüngerer Zeit häufen sich die Erfolgsmeldungen (Abbott 2012), nicht zuletzt dank der massiven Förderung durch die DARPA (Miranda et al. 2015). Diese sah sich durch die steigende Zahl von gliedmaßenamputierten Kriegsveteranen dazu veranlasst, im Jahre 2006 das »Revolutionizing Prosthetics program« zu starten.²⁷ Ziel war die Entwicklung einer anthropomorphen Armprothese, die über alle Eigenschaften eines natürlichen Arms verfügt – also auch über sensorische Kapazitäten und eine vergleichbare Feinmotorik. Das Förderprogramm ist 2010 durch das »Reliable Neural-Interface Technology (RE-NET) program« ergänzt worden, um gleichzeitig den Transfer der invasiven Schnittstellen vom Versuchstier zum Menschen zu beschleunigen.²⁸ Die Bemühungen beginnen offensichtlich weitere Früchte zu tragen: Zwei amerikanische Forschergruppen melden erfolgreiche Implantationen und Ableitungen in Tetraplegikern mit Utah-Arrays (Abb. III.5) – die Patienten beider Gruppen waren in der Lage, komplexe roboterisierte Arme zielgerichtet zu kontrollieren, und das teilweise sogar über 5 Jahre nach der Implantation der Sonden (Collinger et al. 2013; Hochberg et al. 2012). Letzteres ist bemerkenswert, dürfte aber eher eine absolute Ausnahme

²⁷ www.darpa.mil/program/revolutionizing-prosthetics (15.12.2015)

²⁸ www.darpa.mil/program/re-net-reliable-peripheral-interfaces (15.12.2015)

sein, denn eine lange Reihe von Veröffentlichungen konstatiert, wie groß die Problematik der Fremdkörperreaktion für Hirnimplantate ist (Biran et al. 2005; Suner et al. 2005). Diese führt dazu, dass Mikroelektroden vom Körper als Eindringling und Quelle für Entzündungen erkannt und eingekapselt werden. Die so entstehende Glioseschicht verdrängt dann nicht nur die Neuronen von den Ableitkontakten, sondern sie hat auch eine neurotoxische Wirkung und kann zu Neuronenschädigung führen (Silver/Miller 2004). Die Implantate verlieren nach wenigen Wochen ihre Fähigkeit zur vielfachen Ableitung und können dann auch nicht mehr die Signale zur Robotersteuerung liefern – sie werden passiviert. Andererseits kommt es auch immer wieder vor, dass bei Implantaten schon früh die diffizilen Stecker und Zuleitungen aufgrund von Materialermüdung versagen (Sanchez et al. 2006).

Die großen Erwartungen der letzten Dekade, wofür die massive Förderung durch die DARPA exemplarisch steht, scheinen in den letzten Jahren einer gewissen Ernüchterung gewichen zu sein (dazu und zum Folgenden Hofmann 2014, S. 103 f.). Alltagstaugliche Implantate zur Steuerung komplexer Greifprothesen sind derzeit noch Zukunftsmusik. Dennoch sind in den nächsten Jahren weitere große Fortschritte zu erwarten.

Abb. III.5 Steuerung eines Roboterarms mittels implantierter Elektrode



Quelle: nach Hochberg et al. 2012

Es ist damit zu rechnen, dass sich die Datentransferrate der Hirnableitung, welche die Transferrate aus nichtinvasiven BCI jetzt schon beträchtlich übersteigt,



mit noch leistungsfähigeren Elektrodenarrays weiter steigern lässt (Nuyujukian et al. 2014). Dementsprechend wird auch die Bedeutung der Signalverarbeitung zunehmen. Es braucht effizientere und robustere Algorithmen, welche in der Lage sind, aus der wachsenden Datenmenge die Intention des Nutzers zu dekodieren und in Steuersignale umzusetzen. Die Weiterentwicklung sogenannter »intelligenter« Prothesen, die Bewegungen teilautonom unterstützen und damit den Patienten in der Handhabung entlasten, ist in diesem Zusammenhang sicherlich ebenfalls ein wichtiges Forschungsthema für die Zukunft. Allerdings kann auch eine noch so »intelligente« Prothese nicht die sensorische Rückmeldung ersetzen, die für die intuitive und exakte Steuerung einer komplexen Handprothese und insbesondere die Kontrolle des Greifprozesses essenziell ist – die Entwicklung invasiver neuroelektrischer Schnittstellen, die einen bidirektionalen Datentransfer ermöglichen, steht insofern weit oben auf der Forschungsagenda und hat jüngst durch einen europäischen Erfolg einen neuen Schub bekommen (Raspopovic et al. 2014).

Zentrale Voraussetzung für die Bewältigung dieser Herausforderungen sind neben geeigneten Versuchsmodellen vor allem aber langzeitstabile, minimal traumatisierende Implantate (dazu und zum Folgenden Hofmann 2014, S. 103 f.). Der Mangel daran stellt zweifelsohne eines der größten Hindernisse dar, um BMI-Anwendungen aus dem Stadium der Grundlagenforschung in die verbreitete Anwendung zu bringen. Hier sind insbesondere die Materialforschung und Mikrosystemtechnik gefordert, Implantate zu entwickeln, welche die Fremdkörperreaktion und Traumatisierung nach einer Implantation reduzieren, die Liegezeit verlängern und zusätzlich eine bessere und sicherere Ableitqualität durch aktive Sonden und kabellose Datenübertragung bieten. Wie schnell es gelingt (und ob überhaupt), diese Hürden zu überwinden, hängt nicht zuletzt auch vom industriellen Interesse an derartigen Neuroprothesen ab, das gerade in Europa aufgrund der doch ziemlich überschaubaren Zahl an Gliedmaßenamputierten wesentlich geringer als in den USA ausfällt.

Stimulierende Systeme

Im Gegensatz zu den ableitenden Anwendungen befinden sich etliche stimulierende Systeme bereits seit vielen Jahrzehnten im klinischen Einsatz. Dies hat wesentlich damit zu tun, dass es sich vom Grundaufbau her in der Regel um wesentlich einfachere technische Systeme handelt, die insbesondere keine ähnlich hohen Anforderungen an Verfahren der Datengenerierung und -analyse haben. Die stimulierende Wirkung der Elektrizität auf das menschliche Nervensystem wird im Rahmen diverser medizinischer Standardanwendungen sowohl auf invasivem als auch nichtinvasivem Wege ausgenutzt:



- › Zu den ältesten nichtinvasiven Anwendungen gehört der in den 1960er Jahren entwickelte *Fallfußstimulator*. Er wurde entwickelt, um Patienten, die etwa infolge eines Schlaganfalls an der sogenannten Fußheberschwäche leiden, beim Gang zu unterstützen. In der ursprünglichen Version wird der Peroneusnerv im Unterschenkel mittels Klebeelektroden stimuliert, sodass der gesamte Muskel effektiv zur Kontraktion und damit zum Heben des Fußes angeregt wird (Hofmann 2014, S.62 f.). Der exakte Zeitpunkt der Stimulation wird durch einen Fersenschalter ermittelt. Heute gibt es verschiedene kommerzielle kabellose Fußheberstimulatoren, die implantierbare Elektroden und teilweise auch implantierbare Steuerelemente benutzen.
- › Seit den 1950er Jahren wurden diverse invasive *Schrittmacher* für die Wiederherstellung autonomer Funktionen entwickelt, deren Implantation inzwischen zu den medizinischen Routineeingriffen zählt. Herzschrittmacher sind, mit knapp 80.000 neu implantierten Patienten pro Jahr allein in Deutschland,²⁹ mit Abstand am weitesten verbreitet. Zunehmend kommen aber auch Atem- und Blasenschrittmacher zum Einsatz. Zielpersonen sind hier vor allem Querschnittsgelähmte, welche ihre Atmung bzw. Blase nicht mehr kontrollieren können (Hofmann 2014, S. 111 ff.). Anders als bei Herzschrittmachern, die direkt den Herzmuskel stimulieren und diesen so zur Kontraktion anregen, findet hier üblicherweise eine Reizung funktionell relevanter Nerven statt (Atemnerv bzw. Sakralnerv). Anders als die autonom stimulierenden Herz- und Atemschrittmacher werden Blasenschrittmacher durch ein externes Gerät gesteuert, um die Blasenentleerung bei Bedarf auszulösen.
- › Ähnlich wie Herzschrittmacher funktionieren *Vagusnervstimulatoren*, die seit etwa 20 Jahren bei medikamentös schwer behandelbarer Epilepsie und seit Neuestem auch zur Behandlung therapieresistenter Depressionen eingesetzt werden (in Deutschland jedoch derzeit nur in Einzelfällen).³⁰ Der Vagusnerv spielt eine wichtige Rolle im vegetativen Nervensystem. Die Stimulation erfolgt nach der chirurgischen Implantation der Elektroden im Halsbereich und des Schrittmachers in einem voreingestellten Rhythmus³¹ und bewirkt bei Epilepsie eine deutliche Reduktion der Anfallshäufigkeit (Hofmann 2014, S. 109 f.). Die genauen Wirkmechanismen sind noch unbekannt, diverse andere Anwendungsgebiete (Migräne, Demenz, Parkinson etc.) werden derzeit erforscht.

29 <http://de.statista.com/statistik/daten/studie/243619/umfrage/anzahl-der-herzschrittmacher-eingriffe-in-deutschland-nach-art-des-eingriffs/> (15.12.2015)

30 www.psychiatrie.med.uni-goettingen.de/de/content/patienten/239.html (15.12.2015)

31 Zusätzlich gibt es die technische Möglichkeit, die Stimulationshäufigkeit über einen magnetischen Schalter selbst zu beeinflussen.



- › Zur Behandlung chronischer Schmerzen, die anderweitig nicht behandelbar sind, hat sich seit Jahrzehnten die *Rückenmarkstimulation* bewährt. Unterschiedliche Stimulatoren sind kommerziell erhältlich und zugelassen, wobei die einzelnen Elektrodenträger und der Implantationsort die wesentlichen Unterscheidungsmerkmale darstellen (dazu und zum Folgenden Hofmann 2014, S.73 f.). Die Platzierung wird entsprechend dem schmerzenden Körperbereich ausgewählt, sodass die Weiterleitung der Schmerzsignale aus dem Rückenmark ans Gehirn unterdrückt wird und der Patient nur noch ein Kribbeln im vorher schmerzenden Bereich verspürt.

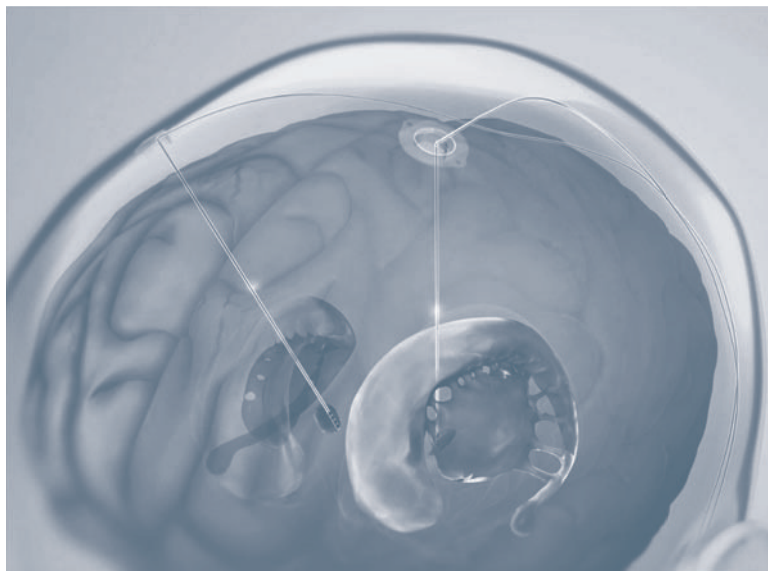
Alle diese Anwendungen sind ausgereift und klinisch erprobt, ihr medizinischer Nutzen steht außer Frage. Ihre Risiken und Nebenwirkungen sind überschaubar, sieht man einmal von operativen Komplikationen ab, die natürlich nie ganz auszuschließen sind. Da die beschriebenen Systeme alle am peripheren (und zudem noch in der Regel am autonomen) Nervensystem andocken, hält sich ihre Brisanz im Sinne der eingangs beschriebenen Problematik der Mensch-Maschine-Entgrenzung jedoch in Grenzen – hauptsächlich, weil sie quasi autonom und technisch außerhalb der Reichweite des Bewusstseins agieren, rein der Funktionswiederherstellung bzw. der Rehabilitation dienen und somit nicht für ein Enhancement menschlicher Fähigkeiten infrage kommen. Anders sieht dies bei zwei weiteren Typen von invasiven Stimulationsanwendungen aus, der tiefen Hirnstimulation und den sensorischen Neuroprothesen, die direkt auf das Gehirn bzw. die Sinneswahrnehmung einwirken und somit ganz neue ethische Fragen aufwerfen.

Tiefe Hirnstimulation

Die tiefe Hirnstimulation (DBS) wird seit 1987 eingesetzt, um neurologische Bewegungsstörungen wie Morbus Parkinson (PD), essenzieller Tremor oder Dystonie symptomatisch zu therapieren (dazu und zum Folgenden Hofmann 2014, S.76 f. u. 117). Das Anwendungsspektrum des Therapieansatzes hat sich in den letzten Jahren sukzessive erweitert, in Europa besteht inzwischen auch eine Zulassung für bestimmte Formen von Zwangsstörungen und Epilepsie – die Eignung für die Behandlung einer ganzen Reihe psychiatrischer Erkrankungen (darunter chronische Depressionen und Tourettesyndrom) wird mit teilweise sehr vielversprechenden Ergebnissen intensiv klinisch erforscht (Bronstein et al. 2011; Laxton et al. 2013; Lozano/Lipsman 2013; Tronnier et al. 2002). Es ist also mit einem wachsenden Markt zu rechnen, wobei schon die bisherigen Fallzahlen beeindruckend sind. Sogenannte »Hirnschrittmacher« wurden inzwischen bei rund 125.000 Patienten weltweit eingesetzt und sind damit in bestimmten Indikationsgebieten schon fast eine Selbstverständlichkeit (Medtronic

2015). Dies ist äußerst bemerkenswert, handelt es sich doch um einen invasiven Therapieansatz: Mehrpolige Elektroden werden in einer stereotaktischen, d.h. minimalinvasiven Prozedur möglichst nahe an die relevanten Gehirnzentren verbracht (meist beidseitig), um dort – verbunden mit einem unter dem Schlüsselbein implantierten Impulsgeber – dauerhaft hochfrequente Impulse abzugeben (Kuhn et al. 2010) (Abb. III.6).

Abb. III.6 Tiefe Hirnstimulation: Stimulation des Nucleus subthalamicus zur Behandlung von Parkinson



Quelle: Medtronic GmbH

Der Erfolg der Therapie hängt einerseits wesentlich von der präzisen Platzierung der Elektroden in der Zielregion ab, die je nach Indikation verschieden ist. Vor der Implantation wird deshalb mittels bildgebender Verfahren ein möglichst genaues Bild des Schädelinneren erzeugt und anschließend durch vorab eingeführte Mikroelektroden eine Teststimulation durchgeführt (bei vollem Bewusstsein des Patienten).³² Der genaue Stimulationsort ist durch die Verwendung mehrpoliger Elektroden aber auch nachträglich noch in sehr engen Grenzen veränderbar (Kuhn et al. 2010). Die Stimulationswirkungen hängen andererseits zentral von den verabreichten Stimulationsparametern ab, die postoperativ oftmals in einem langwierigen

³² www.uniklinik-freiburg.de/neurologie/behandlung/bewegungsstoerungen/tiefe-hirnstimulation.html (15.12.2015)



Prozess feinjustiert werden müssen,³³ um sie der individuellen Symptomatik des Patienten anzupassen.³⁴ Obwohl die genauen Wirkmechanismen noch im Dunkeln liegen, scheint entscheidend zu sein, dass die DBS nicht einfach unspezifischen Einfluss auf das Aktivitätsverhalten aller umgebenden Neuronen nimmt, sondern offenbar nur die neuronale Ausschüttung des Neurotransmitters γ -Aminobuttersäure (GABA) hervorruft (Feuerstein et al. 2011; Hofmann 2014, S. 77). Dieser hat einen hemmenden Einfluss auf überaktive Nervenzellen, wodurch sich die therapeutischen Erfolge – etwa die deutliche Verbesserung der motorischen Kontrolle bei Bewegungsstörungen – zum Teil erklären lassen.

Viele schwer Parkinsonkranke können mithilfe der DBS wieder ein (in mancherlei Hinsicht) beinahe normales Leben führen. Der therapeutische Nutzen gilt deshalb bei den etablierten Indikationen als groß, umso mehr, als die Wirkungen und somit auch mögliche Nebenwirkungen reversibel sind, das heißt nach Abbruch der Stimulation wieder verschwinden. Dennoch wurde die DBS bislang nur in schweren Fällen angewandt, quasi als ultima ratio, sofern also dem Patienten anderweitig nicht mehr zu helfen ist.³⁵ Der Grund hierfür ist, dass die Implantation eines Hirnschrittmachers zwar als relativ sicher und nebenwirkungsarm gilt, aber keinesfalls als risikolos und nebenwirkungsfrei (dazu und zum Folgenden Hofmann 2014, S. 117 ff.). Alleine schon durch die komplizierte neurochirurgische Implantation lassen sich Komplikationen wie Hirnblutungen nie ganz ausschließen. Die Stimulationselektroden selbst sind zwar sehr stabil und weitestgehend vor Kabelbrüchen geschützt, trotzdem kommen Schäden an Kabeln und Steckverbindern vor. Außerdem muss die Batterie des implantierten Pulsgebers etwa alle 3 bis 5 Jahre erneuert werden – was ebenfalls mit einem entsprechenden OP- und Infektionsrisiko einhergeht.³⁶ Die Impulsgeber sind zudem bei Weitem keine kleinen Geräte, sodass sie an ihrem Implantationsort durch die darüber liegende Haut fühlbar sind und lokal zu Problemen führen können (Infektionen, Nekrosen). Obwohl etliche ausgereifte Systeme verschiedener Hersteller auf dem Markt sind, gibt es also noch reichlich Entwicklungs- und Forschungsbedarf in technischer und medizinischer Hin-

33 Der Impulsgeber ist aus diesem Grund von außen programmierbar und kann per Magnetschalter ein- bzw. ausgeschaltet werden.

34 www.uniklinik-freiburg.de/neurologie/behandlung/bewegungsstoerungen/tiefe-hirn-stimulation.html (15.12.2015)

35 Die Ergebnisse der deutsch-französischen Studie »Earlystim« könnten diesbezüglich jedoch ein Umdenken einleiten, denn die Studiendaten zeigen, dass die Neurostimulation offenbar bereits in einem frühen Krankheitsstadium erfolgversprechend eingesetzt werden kann (Schuepbach et al. 2013).

36 Die Batterielaufzeit ist sehr individuell und u. a. abhängig von den programmierten Stimulationsparametern. Alternativ sind auch transkutan aufladbare Systeme am Markt verfügbar, die allerdings einen routinierten Umgang des Patienten mit dem Ladesystem erfordern (persönliche Mitteilung Dr. Holger Storcks, Medtronic GmbH, vom 4.3.2016).



sicht. Dazu gehören u. a. die mikroskopische Verkleinerung der Stimulationselektroden, die bessere und verlässlichere Zielpunkterfassung sowie eine längere Funktionsdauer der Impulsgeber. Aber auch wenn es gelingt, die Technologie weiter zu verbessern, dürften grundsätzliche ethische Vorbehalte gegenüber einer verbreiteten Anwendung der DBS bestehen bleiben. Sie hängen hauptsächlich mit den teilweise schweren psychischen Nebenwirkungen zusammen, die bei Parkinsonpatienten mit Hirnschrittmacher immer wieder beobachtet wurden – dazu gehören Gedächtnis- (1,1–20 %) und Sprechstörungen (10,8–33 %), Depressionen (1,5–25 %), aber auch schwerere Persönlichkeitsveränderungen wie manische (2–28 %) oder aggressive Zustände (2 %) (Clausen 2010a). Ein besonders eindrückliches und gut dokumentiertes Beispiel ist der Fall des kürzlich verstorbenen, bedeutenden Soziologen Helmut Dubiel, der unter schwerem Parkinson litt und seine Erfahrungen mit einem Hirnschrittmacher in einem Buch ausführlich geschildert hat (Kasten).

Die zunehmend durch systematische Untersuchungen – etwa im Rahmen des ELSA-DBS-Projektes³⁷ (Witt 2013, S. 6 f.) – bestätigten Hinweise auf psychische DBS-Effekte machen deutlich, dass die Medizin mit jedem neurotechnologischen Eingriff ins Gehirn, also dem Organ, das unsere Wesenszüge zentral bestimmt, »ethische[s] Hochsicherheitsgebiet« (Quante 2014, S. 2) betritt. Das scheint selbst dann zu gelten, wenn die Prozedur wie bei DBS reversibel ist, da sich durch den Wechsel zwischen den unterschiedlichen Persönlichkeitszuständen – so legen es die Schilderungen von Dubiel nahe – die Entfremdungserfahrungen erst recht verstärken. Im Falle von DBS hat dies nicht nur eine schwierige Nutzen-Risiko-Abwägung zur Folge, sondern aus medizinethischer Sicht wird auch die Frage virulent, inwiefern das Prinzip der informierten Einwilligung überhaupt noch umsetzbar ist (Witt 2013). Denn wenn die Persönlichkeit eines Patienten sich durch den neurotechnologischen Eingriff in wesentlichen Aspekten verändern kann, und damit unter Umständen auch dessen Sicht auf die Operation,³⁸ stellt sich die Frage, auf welcher Grundlage die ethische Legitimierung der Therapie erfolgen kann (Witt 2013, S. 12): Ist hierfür die Sicht des Patienten vor oder nach der Operation entscheidend?

37 <http://geschichte-ethik.uk-koeln.de/de/forschungsstelle-ethik/forschung/elsa-dbs> (15.12.2015)

38 Dies ist kein reines Gedankenspiel. Es gibt tatsächlich Fallberichte, wonach Patienten durch Behandlung mit DBS in manisch-euphorische Zustände verfallen, die sie selber – ganz im Gegensatz zu ihrem Umfeld – als unproblematisch ansehen und in der Folge auch nicht durch eine Anpassung der Stimulationsparameter beheben lassen möchten (Schermer 2011, S. 3).



»Tief im Hirn«: Der Fall Dubiel

In seinem Buch »Tief im Hirn« (2006) beschreibt Helmut Dubiel, wie erheblich sich der bei ihm implantierte Hirnschrittmacher auf sein Bewusstsein und seine Psyche ausgewirkte: So litt er bei eingeschaltetem Schrittmacher unter einem stark eingeschränkten Sprachvermögen, während er bei ausgeschaltetem Schrittmacher mit den Krankheitssymptomen (depressive Zustände, Tremor) zu kämpfen hatte. Das Erstaunliche an seiner Schilderung ist, wie es ihm offenbar möglich war, alleine durch Ein- und Ausschalten des Schrittmachers wie von Geisterhand zwischen diesen unterschiedlichen Zuständen hin- und herzuwechseln. Er schreibt: »Die schlimmste Nebenfolge der Operation bestand in der Störung des Sprachzentrums im Hirn. ... Um ›der Welt‹ nicht offenbar werden zu lassen, dass das Vokabular, dessen ich mich ohne Schwierigkeiten bedienen kann, schrumpft, beschränke ich mich unbewusst auf die Worte, die ich problemlos bewältigen kann, und befürchte, dass meine Artikulationsgabe meinem Denken, der Begriffsbildung und Urteilskraft immer deutlicher hinterherhinkt. ... Ein Licht im Tunnel der Ungewissheit wurde ein Jahr nach der OP sichtbar, weil eine Neurologin in einer anderen Klinik einfach vorschlug, den Schrittmacher versuchshalber abzustellen. Es war, als ob ein Geist aus mir sprach. In derselben Sekunde kehrte meine Stimme zurück, sonor, wohl artikuliert, nur ein wenig heiser. Interessant war, dass nicht nur das Sprechen im technischen Sinne wieder sofort funktionierte, sondern auch meine Verstandestätigkeit und die kognitiven Funktionen – im buchstäblichen Sinne – wieder angeknipst waren.«

Wie nah Technik bereits an Kernbereiche des Menschlichen herangerückt ist, zeigt Dubiels Schilderung der Folgen einer Reaktivierung seines Schrittmachers: »So faszinierend wie erschreckend war vor allem, dass die Depression von mir abfiel, so als sei ein eisernes Band um meine Seele gesprungen. Faszinierend war die Leichtigkeit dieses Vorgangs. Ein Knopfdruck, bestätigt durch ein kaum hörbares digitales Piepsen, unterstützt von einer winzigen Leuchtdiode, öffnete schlagartig den verhangenen Himmel. Freunde, die ich anrief, meinten, ich wäre frisch verliebt, so fröhlich muss ich geklungen haben. Erschreckend und irgendwie demütigend war die Banalität dieses Vorgangs.«

Quelle: Dubiel 2006, S. 125 ff.

Abgesehen von den praktischen Fragen, die sich daran anschließen, etwa nach der Notwendigkeit geeigneter ethischer Leitlinien (Kuhn et al. 2010, S.111), wird anhand dieser Problematik auch der sehr eigentümliche Werkzeugcharakter von Neuroimplantaten wie DBS deutlich. Indem sie quasi mit dem Gehirn

des Patienten verschmelzen und das Potenzial haben, über dessen Hirnprozesse auch dessen Selbst zu beeinflussen, spielen sie einen ungewöhnlich aktiven, ja umgestaltenden Part im Heilungsprozess (Clausen 2010b, S.96 f.). Die Folge ist eine diffuse Entgrenzung von Mensch und Technik – weniger im räumlichen Sinne, sondern vor allem insofern, als der Patient durch die reversible Aktivierung des Hirnschrittmachers gleichzeitig zum »Objekt und Subjekt technischer Kontrolle« (Müller 2010, S.194) wird und damit zunehmend unklar wird, wer eigentlich wen beherrscht: der Mensch die Technik oder die Technik den Menschen. Zwar ist diese Problematik nicht ganz neu – sie stellt sich in ähnlicher Weise auch bei Pharmakotherapien (Sauter/Gerlinger 2012) –, jedoch spitzt sie sich bei den Neurotechnologien durch deren wesentlich spezifischere Wirkungsweise und ihren technischeren Charakter noch einmal deutlich zu. Diese Entwicklung wirft vielfältige philosophisch-anthropologische Fragen nach dem Menschenbild auf, die vor allem unter dem Stichwort »Cyborg« diskutiert werden (Kap. II.2; Müller 2010). Klar ist, dass mit der technischen Veränderbarkeit automatisch auch die Optimierbarkeit des Menschen zum Thema wird, wenngleich die tatsächliche Realisierbarkeit dieser Vision im Fall von DBS aus ethischen und medizinischen Gründen (erhebliche Eingriffsrisiken vs. unklarer Nutzen) noch auf längere Sicht versperrt sein dürfte (Clausen 2012, S.74).

Sensorische Neuroprothesen

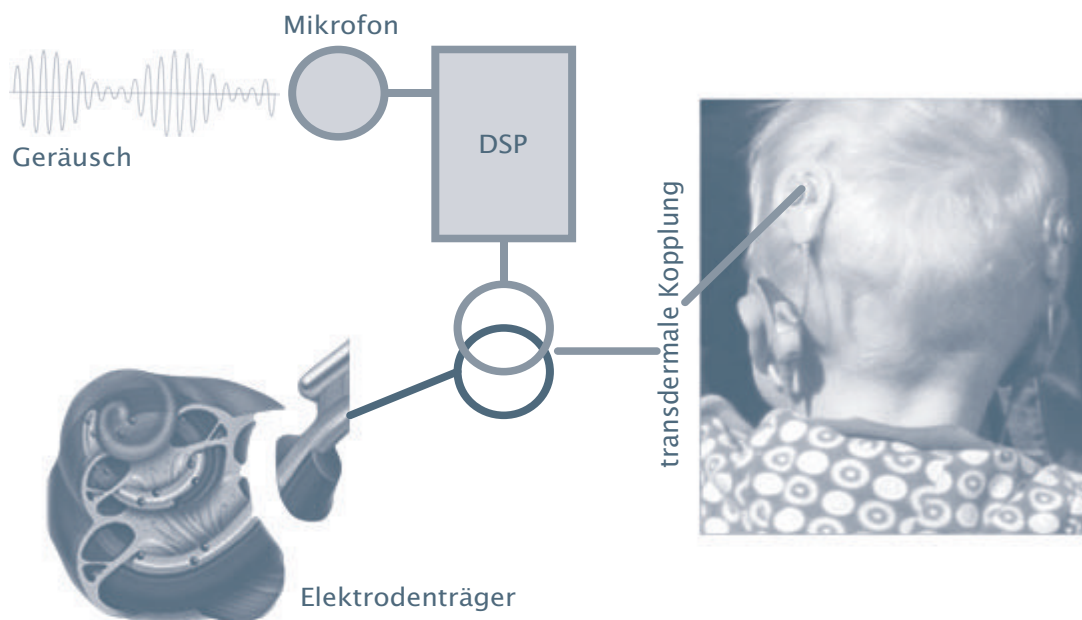
Es ist ein uralte Vision: Blinde wieder sehen, Gehörlose wieder hören zu lassen. Was in biblischen Zeiten noch göttlichen Beistands bedurfte, ist heute dank der modernen Medizin fast schon profane Wirklichkeit geworden – sensorische Neuroprothesen, welche durch die Stimulation peripherer Nerven krankheitsbedingt verloren gegangene Sinnesfunktionen partiell kompensieren, gehören heute zum klinischen Alltag. Die Grundsteine hierfür wurden bereits in den 1950er Jahren gelegt, seither haben vor allem auditorische Neuroprothesen eine beispiellose Erfolgsgeschichte hinter sich, aber auch bei den wesentlich anspruchsvolleren Sehprothesen gibt es erste Durchbrüche zu vermelden.

Zur Wiederherstellung der Hörfunktion kommen zwei unterschiedliche neurotechnologische Herangehensweisen infrage: das *Cochlea-Implantat* (CI) sowie das *auditorische Hirnstammimplantat* (Auditory Brainstem Implant [ABI]) Welcher Therapieansatz der geeignete ist, hängt zentral davon ab, ob der Hörnerv noch funktionsfähig ist. Meistens sind bei Gehörlosigkeit nur die sensiblen Haarzellen geschädigt, die den Schall in elektrische Impulse umwandeln und diese dann an die verbundenen Nervenfasern weitergeben. In solchen Fällen besteht grundsätzlich die Möglichkeit, den intakten Hörnerven mithilfe einer in die Hörschnecke (Cochlea) eingeführten Mikroelektroden-sonde zu stimulieren – so das Grundprinzip des Cochlea-Implantats, das bereits in den

1970er Jahren erstmals klinisch erprobt, 1984 von der US-amerikanischen Food and Drug Agency (FDA) zugelassen und bis heute weltweit über 300.000-mal implantiert wurde (davon 30.000 in Deutschland; dazu und zum Folgenden Hofmann 2014, S.105 f.).³⁹ Moderne CI verfügen über bis zu 22 Kanäle, das heißt einzelne Elektroden, die je nach Platzierung auf dem Hörnerv unterschiedliche Tonwahrnehmungen erzeugen. Neben dem eigentlichen Implantat im Innenohr verfügt ein herkömmliches CI über mehrere außenliegende Komponenten (Abb. III.7).

Abb. III.7

Aufbau eines Cochlea-Implantats



Quelle: nach Hofmann 2014, S. 106

Der Schall wird über ein externes Mikrofon aufgenommen, an einen Signalprozessor weitergeleitet und dort kodiert. Dies ist u. a. deshalb erforderlich, da herkömmliche Stimulationselektroden nur etwa in die ersten eineinhalb der zweieinhalb Cochlea-Windungen eingeführt werden können und so ausschließlich in hohen Frequenzbereichen Höreindrücke erzeugen. Außengeräusche müssen also mittels des Signalprozessors auf diese wahrnehmbaren Frequenzen abgebildet werden, bevor die Signale schließlich mittels einer magnetischen Spulenkupplung durch die Kopfhaut hindurch (transdermal) an das Implantat weiter-

³⁹ <http://schnecke-online.de/informieren/behandlung-und-reha/cochlea-implantat.html> (15.12.2015)



geleitet werden. Außerdem hat der Signalprozessor die Aufgabe, Störgeräusche herauszufiltern und so auch in einer unruhigen Geräuschkulisse Sprachverstehen zu ermöglichen.

Da nur etwa 22 Elektroden bis zu 30.000 Hörrezeptoren gegenüberstehen, ist der entstehende Höreindruck sehr individuell und nicht mit dem beim normalen akustischen Hören vergleichbar. Entsprechend braucht es langes Training, um Geräusche differenzieren und Sprache verstehen zu können. Mit modernen Geräten sind Gehörlose jedoch aufgrund der Plastizität des auditorischen Systems nach einiger Zeit zu erstaunlichen Hörleistungen fähig – Musikhören und Sprachverstehen sind dank hochentwickelter Signalprozessoren in der Regel problemlos möglich. Zu beachten ist allerdings, dass das Implantat bei taub geborenen Erwachsenen nur von geringem Nutzen ist, da die Hörbahn nicht ausgereift und das Gehör damit irreparabel geschädigt ist. Eindrucksvolle Therapieerfolge zeigen sich jedoch bei gehörlos geborenen Kindern. Sie sind zu einem weitgehend normalen Hör- und Spracherwerb in der Lage, sofern das CI beidseitig, vor allem jedoch rechtzeitig, das heißt möglichst in den ersten zwei Lebensjahren, eingesetzt wird. Da Gehörlosigkeit medizinisch als Behinderung gilt, haben gehörlos geborene Kinder in Deutschland einen rechtlichen Anspruch auf ein krankenkassenfinanziertes CI (Müller/Zaracko 2010, S.247). Interessanterweise kritisieren aber gerade Vertreter der Gehörlosengemeinschaft diese Praxis teils vehement. Sie befürchten, dass die reiche Gehörlosenkultur mit ihrer eigenen Gebärdensprache dadurch verloren gehen könnte – zumindest jedoch zur defizitären, aus der Not geborenen Behindertenkultur gestempelt wird, die im Grunde genommen über keine eigenständige Existenzberechtigung verfügt (Heffley 2010; Park 2014). Für viele Betroffene ist die Gehörlosenkultur jedoch ein zentraler Bestandteil ihrer Identität. Als Ausweg aus diesem ethischen Dilemma bietet sich an, dass man gehörlosen Kindern – die ja noch nicht autonom entscheiden können, in welcher der »Welten« sie später leben möchten – sowohl eine Hörprothese implantiert als sie auch die Gebärdensprache erlernen lässt (Müller/Zaracko 2010, S.248).

Bei den wenigen gehörlosen Patienten, die über keinen funktionsfähigen Hörnerv mehr verfügen, was häufig aufgrund einer Tumorerkrankung der Fall ist (Neurofibromatose Typ II, [NF2]), versagt auch ein CI (Hofmann 2014, S.106). In solch seltenen Fällen helfen derzeit nur auditorische Hirnstammimplantate weiter, die vom Grundaufbau und der Funktionsweise einem CI analog sind. Der einzige Unterschied ist das Implantat selber, das aus einer Oberflächenelektrode besteht und direkt den am Hirnstamm verbliebenen Hörnervkern (Nucleus cochlearis) elektrisch stimuliert – es setzt also einen Schritt später in der Signalübertragung an und überbrückt damit das gesamte Innenohr. Aus diesem Grund sind die Hörergebnisse auch weniger vielversprechend als bei einem CI und zudem oft nur nach jahrelangem Training erzielbar

(Colletti et al. 2012, S. 355). Patienten mit Neurofibromatose Typ II, für die ABI ursprünglich entwickelt wurden, können immerhin Umgebungsgeräusche wahrnehmen und besser Lippenlesen, erreichen aber in der Regel nur ein rudimentäres Sprachverständnis (Siegbahn et al. 2014). Neuere Erkenntnisse deuten jedoch darauf hin, dass in anderen Indikationsgebieten, für die ABI inzwischen zugelassen wurden (Kinder mit Gehörnervaplasie, postmeningitische Ertaubung bei Erwachsenen etc.), deutlich bessere Ergebnisse erzielt werden können und Patienten insbesondere auch zu einem teilweise exzellenten Sprachverstehen in der Lage sind (Colletti et al. 2012). Dies trifft vor allem auch auf die mehr als 100 Kinder zu, die weltweit mit einem ABI versorgt worden sind. Insgesamt scheint die Variabilität in der Hörqualität jedoch deutlich größer zu sein als bei einem CI, was nicht zuletzt auch damit zu tun haben könnte, dass der chirurgische Eingriff einerseits wesentlich komplizierter ist, andererseits aber auch deutlich seltener vorgenommen wird.

Erklärtes Ziel der Hersteller von Hörprothesen ist die Erweiterung der Indikationsstellung (dazu und zum Folgenden Hofmann 2014, S. 106 ff.). Beim Cochlea-Implantat geht die Entwicklung hin zu hybriden Systemen, die konventionelles Hörgerät und Hörprothese in einem System vereinen und so das Resthörvermögen (besonders in den tiefen Frequenzbereichen) besser erhalten sollen.⁴⁰ Diese Geräte benötigen jedoch aufgrund der unterschiedlichen Signalquellen wesentlich ausgefeiltere Sprachprozessoren. Außerdem soll die bereits sehr gute Spracherkennung weiter verbessert werden, was leistungsfähigere Elektroden-Arrays und mehr Stimulationskontakte erfordert. Insgesamt ist die Entwicklung der elektronischen Hörhilfen mit dem derzeitigen Stand sicherlich noch nicht abgeschlossen.

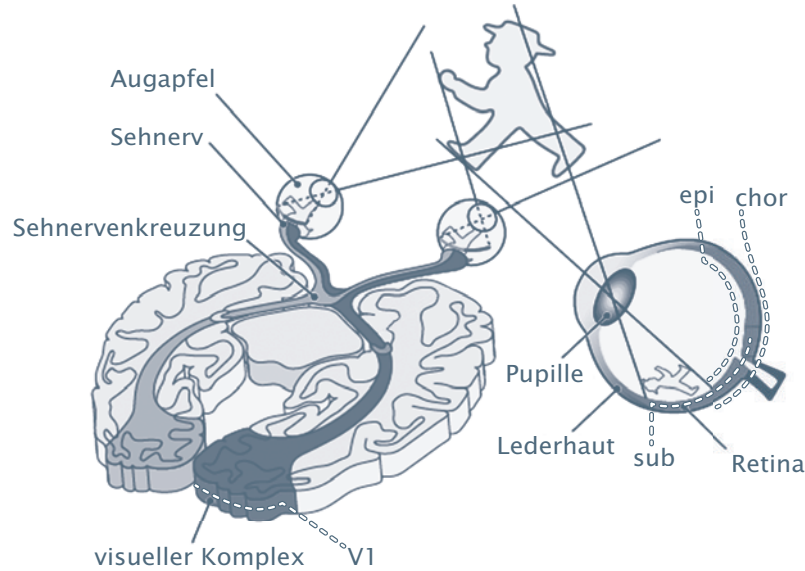
Aufgrund der wesentlich größeren Komplexität des visuellen Systems – 130 Mio. Photorezeptoren stehen 30.000 Hörrezeptoren gegenüber – ist die Entwicklung von *Sehprothesen* deutlich anspruchsvoller (dazu und zum Folgenden Hofmann 2014, S. 108). Abbildung III.8 zeigt eine schematische Darstellung des visuellen Systems mit den möglichen Ansatzpunkten für eine neurotechnologische Stimulation (gestrichelte Bereiche):

- > Die Stimulation des visuellen Kortex bzw. des Sehnervs kommt infrage, wenn der Sehnerv oder die Sehbahn durch Hirntumore oder Unfall geschädigt sind (V1) (Duret et al. 2005).
- > Ist der Sehnerv noch intakt, die Retina jedoch beschädigt, bietet sich eine Stimulation der dortigen Ganglionzellen an verschiedenen Positionen am Augapfel an.

⁴⁰ Abhängig vom Schadensort sind unterschiedliche Implantate möglich bzw. verfügbar: www.hoerzentrum-hannover.de/index.php?id=64 (15.12.2015).

Abb. III.8

Ansatzpunkte für Sehprothesen



Quelle: nach Stieglitz et al. 2014, S. 456

Die ursprüngliche, bereits aus den 1960er Jahren stammende Idee, durch direkte Reizung des visuellen Kortex Seheindrücke zu erzeugen, hat sich bis heute als noch nicht gangbar herausgestellt (dazu und zum Folgenden Hofmann 2014, S. 107 f.).⁴¹ Auch der Versuch, durch Mikrostimulation des primären visuellen Kortex mit dem Utah-Array bessere Ergebnisse zu erreichen, muss wohl als gescheitert gelten (Normann et al. 1999). Ein wenig weiter in der Entwicklung, aber auch noch weit entfernt von jeder klinischen Nutzung ist eine Sehprothese, die das Utah-Array in den Sehnerv implantieren und damit stimulieren möchte (Duret et al. 2005). Auf dem besten Wege, die klinische Erfolgsgeschichte des Cochlea-Implantats zu wiederholen, sind hingegen sogenannte *Retina-Implantate*, welche die verbleibenden Nervenzellen der Retina stimulieren. Sie wurden primär zur symptomatischen Behandlung von degenerativen Netzhauterkrankungen

41 Trotzdem gelang es William Dobbelle Ende der 1990er Jahre, ein beachtliches Presseecho zu erhalten, als er einem blinden Erwachsenen Stimulationselektroden auf den visuellen Kortex implantierte und mit einem kameragebundenen Stimulationssystem verband (Hofmann 2014, S. 107). Angeblich war der Blinde damit sogar in der Lage, ein Auto über einen leeren Parkplatz zu lenken (Kotler 2002). Eine klinische Anwendung fand sich für dieses System jedoch nicht und das Dobbelle-Institut, das aus rechtlichen Gründen nach Portugal verlegt worden war, beendete seine Arbeit nach dem Tod des Gründers 2004.



kungen wie der genetisch bedingten Retinitis Pigmentosa entwickelt, an der in Deutschland etwa 40.000 Menschen leiden.⁴²

Grundidee der Retina-Implantate ist es, mikrotechnisch erzeugte, flache Elektrodenarrays an eine Stelle auf (epiretinal), unter (subretinal) oder hinter der Netzhaut (suprachoroidal) so zu platzieren, dass die elektrische Stimulation der mikroskopischen Kontaktgitter die Aktivität der zerstörten Photorezeptoren ersetzt (dazu und zum Folgenden Hofmann 2014, S. 107 f.). Die Platzierung der stecknadelgroßen Elektrodenarrays erfolgt durch eine kleine Öffnung in der Lederhaut des Augapfels und durch geschickte intraoperative Platzierung. Bei den meisten Systemen wird die Arrayfolie kabellos oder durch einen Stecker mit dem Bildgewinnungs- und Verarbeitungssystem verbunden, und die Stimulation der einzelnen Kontaktpunkte erfolgt dann entsprechend von außen. So ist es auch bei dem weltweit ersten zugelassenen Retina-Implantat der Fall – »Argus II« von Second Sight –, mit dem angeblich 5 bis 10% Sehvermögen erreicht werden sollen (Humayun et al. 2012). Wesentliche Bestandteile des Systems werden extern am Körper getragen: Das Bildsignal wird über eine winzige, an einer Spezialbrille befestigte Kamera aufgenommen, dann an einen Minicomputer übermittelt, kodiert und schließlich kabellos an das eigentliche Implantat übermittelt. Fast gänzlich ohne externe Hardware und damit besonders innovativ ist das Alpha-IMS-System der Retina Implant AG, einer Ausgründung der Universität Tübingen. Die Gewinnung eines Abbildes der Außenwelt erfolgt mit den Photodioden eines implantierten, hochauflösenden Mikrochips, der über eine externe Stromquelle transdermal mit Energie versorgt wird. Die Ergebnisse sind vielversprechend: Blinde Patienten können Gegenstände wieder erkennen und teils sogar Gesichtsmimik und Buchstaben lesen (Stingl et al. 2013). Das seit 2013 ebenfalls CE-zertifizierte Alpha-IMS-System befindet sich in einer frühen Phase der Markteinführung, weltweit wurden bislang etwas mehr als 40 Patienten damit behandelt; in Deutschland werden die Kosten für die entsprechende Behandlung der Retinitis Pigmentosa seit 2014 von den Krankenkassen übernommen.⁴³

Cyborgs: Vision oder bereits Realität?

Im Gegensatz zu den Cochlea-Implantaten handelt es sich bei den wenigen, kommerziell erhältlichen Retina-Implantaten noch um teure Nischenprodukte, die bislang nur für die symptomatische Therapie der relativ seltenen Erbkrank-

42 www.pro-retina.de/netzhauterkrankungen/retinitis-pigmentosa/krankheitsbild (15.12.2015)

43 www.aerzteblatt.de/nachrichten/59712/Retina-Implantat-Kostenerstattung-durch-gesetzliche-Krankenkassen (15.12.2015)



heit Retinitis Pigmentosa zugelassen sind. Mit der weiteren Verbesserung der Systeme, insbesondere was die Auflösung der Mikrochips betrifft, ist jedoch in Zukunft mit einem erweiterten Indikationsbereich und einer Zulassung etwa auch für die altersbedingte Makuladegeneration zu rechnen, von der in Deutschland mehr als eine Million Menschen betroffen sind.⁴⁴

Unabhängig von medizinischen Nutzen- und Anwendungsfragen lässt sich insbesondere am Beispiel der zunehmend alltäglicher und leistungsfähiger werdenden Sinnesprothesen (und auch in begrenzterem Umfang in Zusammenhang mit der tiefen Hirnstimulation) symptomatisch verfolgen, wie Mensch und Maschine zunehmend zu einer Einheit verwachsen – und sich damit herkömmliche Vorstellungen einer körperlichen Norm aufzulösen beginnen. Denn die mit Cochlea- und Retina-Implantaten ermöglichte Sinneserfahrung hat zwar einerseits einen rein therapeutischen Zweck, dient also der Wiederherstellung der Hör- oder Sehfunktion, ist aber andererseits zu einem Grade künstlich, der eine klare Abgrenzung zwischen technischem Hilfsmittel und körperlicher Funktion ad absurdum führt. Beim Cochlea-Implantat handelt es sich ja nicht um eine bloße *Hörhilfe*, sondern um einen elektronischen *Hörersatz*. Das heißt, der Träger einer Hörprothese hört anders als der Normalhörende, und er muss das Hören ganz neu erlernen. Letzteres hat auch ganz konkrete Auswirkungen auf neuronaler Ebene, insofern nämlich, als sich auch das Gehirn plastisch auf die neuen Signale einstellen muss, um sie verarbeiten zu können, und so mit der Neuroprothese quasi zu einer Funktionseinheit verschmilzt.

Vor diesem Hintergrund erscheint, auch angesichts der großen Menge an Personen mit Cochlea-Implantat (weltweit über 300.000 Personen, davon ca. 30.000 in Deutschland), die Frage berechtigt, ob die insbesondere auch in den Diskussionen über Transhumanismus und HE weitverbreiteten Cyborgvisionen – also die Verschmelzung von Mensch und Technik zu einer Art kybernetischem Mischwesen (Kap. II.2) – möglicherweise bereits heute schon Realität geworden sind (Service 2013). Explizit bejaht wird dies von einer wachsenden Zahl von Aktivisten, die sich als Cyborgs bezeichnen und sich auch in Deutschland in Plattformen wie dem 2013 gegründeten Cyborg e. V., einer »Gesellschaft zur Förderung und kritischen Begleitung der Verschmelzung von Mensch und Technik«⁴⁵, zu vernetzen beginnen. Eine Kernforderung dieser Gesellschaft ist das Recht auf die technologische Modifikation des eigenen Körpers. Im Gegensatz zum Hauptstrom des transhumanistischen Diskurses verzichtet der Cyborg e. V. aber darauf, als übergeordnetes Ziel der Körpermodifikation eine Steigerung der menschlichen Leistungsfähigkeit, eine Verbesserung des Menschen oder gar eine Überwindung menschlicher Körperlichkeit zu nennen. Diese Herangehens-

44 www.medical-tribune.de/medizin/fokus-medizin/artikeldetail/wie-moderne-technik-blinde-sehen-laesst.html (15.12.2015)

45 <https://cyborgs.cc/> (15.12.2015)

weise, die sich vom vorrangig melioristisch denkenden Transhumanismus abgrenzt, beinhaltet auch die Ablehnung einer jeglichen »Normierung des Menschen«.46

Während neurotechnologische Optimierungsmöglichkeiten für Nichtbehandlungsbedürftige aus Kosten- und Gesundheitsgründen auf längere Sicht verstellt sein dürften und tatsächlich in den Bereich des Visionären gehören, eröffnen sich Trägern von Neuroprothesen wie Enno Park, dem Gründer von Cyborg e. V., ganz neue Optionen. So bietet das von ihm genutzte Cochlea-Implantat bereits die Möglichkeit, Störgeräusche auch in sehr lauter Umgebung (z. B. bei einem Konzert) quasi auszublenden, womit er imstande ist, in solchen Situationen z. T. besser zu hören als ein Nichthörgeschädigter. Im Prinzip wäre es zudem bereits heute ohne Weiteres durchführbar, den Sprachprozessor eines Cochlea-Implantats so umzuprogrammieren, dass sinneserweiternde Funktionen ermöglicht werden – beispielsweise das Hören von Ultraschall. Rein aus rechtlicher Sicht spräche hier wohl nichts dagegen, praktisch lässt es sich derzeit jedoch kaum umsetzen, da Schnittstellen und Software u. a. aus patentrechtlichen, aber auch Sicherheitsgründen nicht offengelegt und damit nicht frei konfigurierbar sind. Offene Standards zu etablieren, gehört daher zu den primären Zielen des Cyborg e. V., schließlich gehe es dabei um die Selbstbestimmung über den eigenen Körper.47 Und Selbstbestimmung bedeutet hier in erster Linie Selbsttechnisierung, im Sinne der technologischen Modifizierung des eigenen Körpers.

Interessant ist in diesem Zusammenhang auch der Fall des britischen Künstlers und Cyborgaktivisten Neil Harbisson, der an einer angeborenen Farbenblindheit leidet. Zur Kompensation dieser Einschränkung trägt er einen sogenannten Eyeborg (Abb. III.9), eine per Knochenimplantat fest mit dem Schädel verbundene Antennenvorrichtung, die es erlaubt, Farben akustisch wahrzunehmen (Pearlman 2015). Harbisson gilt nicht nur als der erste offiziell anerkannte Cyborg – die britischen Behörden erlaubten ihm nach einigen Hin und Her, sich mit Eyeborg auf dem Passfoto ablichten zu lassen –, sein Beispiel macht zudem noch einmal deutlich, wie die Grenzen zwischen Therapie und Enhancement durch die neuen technologischen Optionen zunehmend verschwimmen. Nicht nur, dass er Farben hören kann, sein elektronisches »Auge« bietet ihm darüber hinaus die Möglichkeit, auch in Frequenzbereiche vorzudringen (Ultraviolett und Infrarot), die der natürlichen Wahrnehmung verwehrt sind. Das Gerät wird mittlerweile testweise auch von Blinden verwendet.

Bei Fällen, wie dem von Neil Harbisson, handelt es sich aber noch um Einzelbeispiele. Auch die Zahl der Menschen, die aus medizinischen Gründen oder

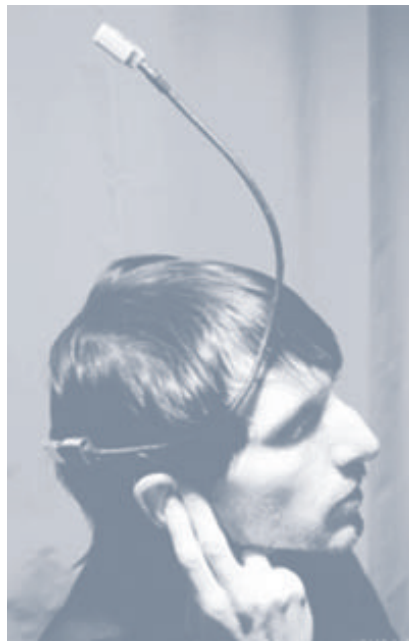
46 <http://wiki.cyborgs.cc/wiki/Manifest> (15.12.2015)

47 www.zeit.de/digital/internet/2013-07/sigint-enno-park-german-cyborg-society/komplettansicht (15.12.2015)

zur Behinderungskompensation Neurotechnologien nutzen und sich aus diesem Grund als Cyborgs definiert haben – sei es wie Enno Park im positiven oder wie der bereits erwähnte Helmut Dubiel im negativen Sinne (als »Cyborg wider Willen«⁴⁸) –, ist weiterhin gering.

Abb. III.9

Neil Harbisson: Cyborg mit Eyeborg



Quelle: nach Moon Ribas/Wikimedia Commons/CC-BY-3.0

Vor allem aber sind Sinn und Zweck der bereits realisierten bzw. angestrebten neurotechnologischen *Enhancementmöglichkeiten* äußerst fraglich. Welcher Nutzen besteht beispielsweise darin, Ultraschall oder Infrarot wahrzunehmen? Geht damit tatsächlich eine (wie auch immer definierte) *Verbesserung* des menschlichen Körpers einher? Unter den aktuellen sozialen und technischen Rahmenbedingungen jedenfalls ist bei nüchterner Betrachtung kein nennenswerter Zusatznutzen von gesellschaftlicher oder auch nur individueller Relevanz erkennbar, keiner zumindest, der eine Operation bei einem körperlich nicht beeinträchtigten Menschen rechtfertigen könnte. Zudem ist in Bezug auf Cochlea-Implantate darauf hinzuweisen – was auch Enno Park vielfach getan hat –, dass den wenigen existierenden oder zukünftig möglichen Zusatzfähigkeiten (wie das überdurchschnittliche Hören in sehr lauten Umgebungen)

48 www.wiwo.de/technologie/forschung/helmut-dubiel-ich-kam-mir-vor-wie-ein-versuchskaninchen/11336310.html (15.12.2015)

weiterhin unterdurchschnittliche Hörfähigkeiten in vielen Situationen gegenüberstehen.

Forderungen und Praktiken der Cyborgisierung sind aktuell vor allem als ein Spiel mit technologischen Möglichkeiten und den durch sie beflügelten Zukunftsvisionen zu werten. Das heißt allerdings nicht, dass die Debatten zum Cyborgkonzept – oder gar die Diskussionen über als Cyborgisierungsprozesse begreifbare wissenschaftlich-technische Entwicklungen – müßig oder verfrüht wären. Angesichts der neurotechnologischen Entwicklungsdynamik erscheint eine möglichst frühzeitige gesellschaftliche Auseinandersetzung mit den ethischen, rechtlichen und gesellschaftlichen Implikationen dieser neuen Technisierungsprozesse – und insbesondere auch eine Verständigung über das gesellschaftlich Wünschbare – durchaus als angebracht (Müller 2010, S. 189).

Forschungslandschaft in Deutschland

1.3

Die gesamte Forschungs- und Entwicklungslandschaft im Umfeld der beschriebenen Neurotechniken gründet in der Bundesrepublik Deutschland auf einer weltbekannten und sehr starken neurowissenschaftlichen Grundlagenforschung (dazu und zum Folgenden Hofmann 2014, S. 103 u. 127 ff.). Deutsche Forschungsstandorte verfügen im Bereich der neuronalen Schnittstellen in der Regel über eine hervorragende Expertise, sie scheitern aber häufig daran, die Nützlichkeit neurotechnischer Entwicklungen am Primaten oder klinisch zu zeigen.

Insbesondere die BCI-Forschung ist hierzulande prominent vertreten – als Zentrum der deutschen BCI-Forschung gilt das Institut für Medizinische Psychologie und Verhaltensneurobiologie am Universitätsklinikum Tübingen⁴⁹ mit der Gruppe um Niels Birbaumer sowie das dortige Max-Planck-Institut (MPI) für biologische Kybernetik⁵⁰, das sich mit der Weiterentwicklung bildgebender Verfahren beschäftigt (dazu und zum Folgenden Hofmann 2014, S. 128). In den letzten Jahren ist generell und auch in Deutschland eine deutliche Zunahme der Forschungstätigkeit im BCI-Feld zu beobachten. Neben den klassischen Disziplinen, die sich traditionell mit EEG-Signalen beschäftigen, also Neurologie und Psychologie, haben die entsprechenden Datensätze auch die Aufmerksamkeit der Machine-Learning-Community geweckt, welche die Algorithmen zur Analyse dieser hochkomplexen Daten weiterzuentwickeln sucht. Diesbezüglich kommen wichtige Impulse aus dem BMBF-geförderten

49 www.medizin.uni-tuebingen.de/Zuweiser/Institute/Medizinische+Psychologie.html (15.12.2015)

50 www.kyb.tuebingen.mpg.de/de.html (15.12.2015)



»Berlin Brain-Computer Interface«⁵¹ an der TU Berlin. Auch Gruppen am Deutschen Forschungszentrum für Künstliche Intelligenz⁵², an der European Medical School in Oldenburg⁵³ ebenso wie am Forschungszentrum Jülich⁵⁴ sind hier von Bedeutung.

Die deutsche BMI-Forschung ist im Vergleich dazu international weniger sichtbar (dazu und zum Folgenden Hofmann 2014, S.128 f.). Nur vereinzelte Standorte, hauptsächlich im Umfeld der Primatenforschung (Bremen, Göttingen und Tübingen) oder der Epilepsiechirurgie (Freiburg und München), beschäftigen sich dezidiert mit der invasiven Gewinnung von elektrophysiologischen Daten aus dem Gehirn von Tiermodellen. So gibt es zwar eine ganze Reihe von erstklassigen neurophysiologischen Gruppen, aber diese arbeiten meist am rein neurowissenschaftlichen Erkenntnisgewinn – weniger an der konkreten Anwendung. Als wichtige und herausragende Ausnahme ist im Rahmen der Forschung zum künstlichen Sehen die Universitäts-Augenklinik Tübingen⁵⁵ zu nennen, die zusammen mit der Universität Marburg die ersten deutschen Anträge für ein Retina-Implantat gestellt hatte. Inzwischen wurde ein Teil des Know-hows von kommerziellen Herstellern übernommen, jedoch gibt die Tübinger Augenklinik immer noch das Maß vor, an dem sich die Forschung orientiert. Bezüglich anderer stärker klinisch ausgerichteter Forschungstätigkeiten in Deutschland sind folgende Fakten noch erwähnenswert:

- › Cochlea-Implantate werden nach Herstellerangaben an 85 Standorten in Deutschland implantiert. Die Technologie wird an der Medizinischen Hochschule Hannover, der Universität Köln und der Universität Oldenburg weiterentwickelt.
- › Die Implantation eines Hirnschrittmachers durch eine stereotaktische Operation wird inzwischen an über 20 Standorten bundesweit angeboten. Deutschlandweit finden auch wissenschaftliche Aktivitäten statt, die von der Arbeitsgemeinschaft »Tiefe Hirnstimulation« koordiniert werden.⁵⁶
- › Maßgebliche Forschungen zur funktionellen Elektrostimulation finden sich u. a. an der Klinik für Paraplegiologie des Universitätsklinikums Heidelberg⁵⁷, während die transkranielle Gleichstromstimulation (tDCS) an der Klinik für Klinische Neurophysiologie der Universitätsmedizin Göttingen⁵⁸ mitentwickelt wurde und dort weiterhin intensiv und weltweit

51 www.bbci.de/ (15.12.2015)

52 www.dfki.de/web (15.12.2015)

53 www.uni-oldenburg.de/european-medical-school/ (15.12.2015)

54 www.fz-juelich.de/portal/DE/Home/home_node.html (15.12.2015)

55 www.medizin.uni-tuebingen.de/augenklinik/ (15.12.2015)

56 www.tiefehirnstimulation.de/ (15.12.2015)

57 www.klinikum.uni-heidelberg.de/Willkommen.115090.0.html (15.12.2015)

58 www.neurologie.uni-goettingen.de/index.php/startseite.html (15.12.2015)



führend untersucht wird. Entwicklungen zur Neuroprothetik werden an diversen Fraunhofer-Instituten vorangetrieben, hier ist vor allem das Fraunhofer-Institut für Biomedizinische Technik⁵⁹ in St. Ingbert hervorzuheben.

Trotz einzelner herausragender Gegenbeispiele hat die deutsche FuE-Landschaft – vor allem in Relation zu ihrem hervorragenden akademischen Ruf – insgesamt nicht allzu viele Entwicklungserfolge vorzuweisen (dazu und zum Folgenden Hofmann 2014, S.129 f.). Dazu bedürfte es auf nationaler Ebene u. a. einer gezielteren und risikobereiteren Forschungsförderung. Was hierzu ebenfalls fehlt, ist ein Zentrum für Neurotechnik und Neuroprothetik, wie es in der Schweiz (Wyss Center for Neuroprosthetics),⁶⁰ in Belgien (Neuroelectronics Research Flanders),⁶¹ Dänemark (SMI Aalborg)⁶² oder den USA (Pittsburgh, Brown, Duke) zu finden ist. Ein interdisziplinäres Forschungszentrum auf höchstem akademischem Niveau könnte als eine Art Kondensationskeim fungieren und dabei helfen, die akademisch hochrangige deutsche Neuroforschung zu bündeln, Ausgründungen zu ermöglichen und Industriepartner anzulocken.

Fazit und Ausblick

1.4

Wie der technologische Überblick gezeigt hat, befinden sich die meisten der zuvor beschriebenen Anwendungen noch in frühen Phasen der experimentellen Erprobung, mit wenigen Ausnahmen (Cochlea-Implantat, tiefe Hirnstimulation) sind sie auf dem Stand von erweiterten Machbarkeitsbeweisen (Proofs of Principle) (Hofmann 2014, S.131). Grundsätzlich gilt, dass sich die Entwicklung nützlicher und verlässlicher neurotechnologischer Anwendungen vor eine Reihe prinzipieller Hürden gestellt sieht, deren Art und Ausprägung wesentlich vom Grad der Invasivität der zugrundeliegenden neuroelektrischen Schnittstellen abhängt (unabhängig von der Signalrichtung). Nichtinvasive Schnittstellen auf der einen Seite leiden an unspezifischen, räumlich schlecht aufgelösten Signalen (EEG) oder hohem apparativem Aufwand (fMRI), was ihre Nutzbarkeit außerhalb des Labors stark einschränkt. Invasive Schnittstellen auf der anderen Seite ermöglichen zwar den zielgenauen Austausch höherer Datenraten, machen dafür aber einen chirurgischen Eingriff erforderlich mit entsprechenden gesundheitlichen Risiken – mehr noch, der Körper rea-

59 www.ibmt.fraunhofer.de/ (15.12.2015)

60 <http://cnp.epfl.ch/> (15.12.2015)

61 www.nerf.be/ (15.12.2015)

62 www.smi.hst.aau.dk/ (15.12.2015)



giert auf alle Implantate mit einer starken Fremdkörperreaktion und limitiert dadurch deren Verträglichkeit bzw. Liegezeit und damit auch den langfristigen Nutzen darauf basierender Anwendungen (Hofmann 2014, S.5). Trotzdem wird bereits eine ganze Reihe der besprochenen neuroelektrischen Schnittstellen schon mehr oder weniger erfolgreich klinisch angewendet und ist damit grundsätzlich geeignet, im Rahmen von kommerzieller Forschung weiterentwickelt zu werden (Hofmann 2014, S.89). Aufgrund ihrer deutlich höheren Bandbreite und der besseren Datenqualität verfügt die invasive Herangehensweise dabei anscheinend prinzipiell über das größere Zukunftspotenzial, insbesondere im Hinblick auf die angestrebte Interaktion mit komplexen Maschinen (Hofmann 2014, S.131; del R. Millán/Carmena 2010). Die Anwendungsperspektiven nichtinvasiver Schnittstellen sind zwar diesbezüglich eher begrenzt, dennoch ist davon auszugehen, dass sie aufgrund ihrer relativ sicheren und risikolosen Anwendbarkeit perspektivisch eine weit größere Verbreitung erfahren dürften – beispielsweise als Kommunikationshilfe für Schwerstgelähmte, als therapeutisches Hilfsmittel (Neurofeedback und Rehabilitation; Birbaumer/Chaudhary 2015) oder als Unterstützungstechnologie in der Pflege. Vor dem Hintergrund der begrenzten Bandbreite nichtinvasiver Mensch-Maschine-Schnittstellen erweist sich die zunehmende Autonomisierung der maschinellen Regelung (etwa in Form »intelligenter« Prothesen) als vielversprechende und intensiv beforschte Option, um mit den eingeschränkten Steuerungsmöglichkeiten solcher Systeme umgehen zu können (sogenannte »shared autonomy«). BCI-Anwendungen stellen folglich ein konkretes Feld dar, in dem zukünftig mit einer zunehmenden Konvergenz zwischen Robotik und Neurotechnologien zu rechnen ist.

Im Rahmen der invasiven Interaktion mit dem Nervensystem ist in jüngster Zeit eine neue Anwendungsdimension am Horizont aufgetaucht, die mit der Entwicklung bidirektionaler Schnittstellen einhergeht. Ziel ist die Entwicklung naturnaher Regelkreise durch Integration ableitender und stimulierender Verfahren. Konkret stehen zwei Anwendungsgebiete im Vordergrund, nämlich die Entwicklung von »fühlenden« Prothesen sowie von aktiven Implantaten (auch Elektrozeutika genannt) (dazu und zum Folgenden Hofmann 2014, S.120 ff.):

- › *Fühlende Prothesen:* Herkömmliche Prothesen geben dem Träger höchstens ein minimales Feedback, das in der Regel nur als grober Widerstand wahrnehmbar ist. Das feinfühliges Greifen z. B. kleiner oder zerbrechlicher Gegenstände ist so nicht möglich, zudem wird die Prothese unter diesen Umständen als Fremdkörper wahrgenommen. Mit der fortlaufenden Entwicklung immer komplexerer Handprothesen, die anspruchsvolle Bewegungsfunktionen bieten, steigt deshalb auch der Bedarf an Möglichkeiten sensorischer Rückmeldung über bidirektionale Schnittstellen. Ansatzweise lässt sich dies offenbar bereits mit dem TMR-Verfahren realisieren, da die



in den Brustraum verpflanzten Armnerven bei Berührung auch sensorische Reize an das Gehirn übermitteln – inwiefern sich so eine sensorische Rückmeldung aus der Prothese realisieren lässt, wird derzeit untersucht (Graimann 2015, S.25). Eine direktere Integration von Sensorik und Aktorik verspricht der invasive Weg, der zurzeit intensiv erforscht wird. Mittels einer kurzzeitig implantierten Nervenschnittstelle gelang es jüngst im Rahmen des europäischen NEBIAS-Projekts⁶³, an dem auch deutsche Wissenschaftler beteiligt sind, nicht nur Steuersignale aus dem Armstumpf zu extrahieren, sondern auch Drucksignale aus der Prothese in die Nerven zurückzusenden (Raspopovic et al. 2014).⁶⁴ Die Versuchsperson war also nicht nur in der Lage, die Unterarmprothese zu steuern, sondern mit ihr auch etwas zu fühlen. Dieser erste Prototyp demonstriert nicht nur die prinzipielle Machbarkeit einer Closed-Loop-Prothese, sondern zeigt auch, dass sich neben dem Gehirn möglicherweise auch das periphere Nervensystem erfolgversprechend für den Anschluss einer Armprothese nutzen lässt.

- › *Elektrozeutika*: Praktisch alle aktuell erhältlichen Schrittmacher arbeiten kontinuierlich innerhalb der vom betreuenden Arzt eingestellten Parameter und Settings. Es gibt jedoch Hinweise, dass beispielsweise Hirnschrittmacher unter Umständen effektiver und schonender eingesetzt werden können, wenn die Stimulation bedarfsgesteuert erfolgt. Vor Kurzem wurde in den USA erstmals ein System zur Unterdrückung epileptischer Anfälle zugelassen, das diesem Prinzip folgt. Es überwacht automatisch die Hirnaktivität über eine ableitende Tiefenelektrode und wird nur bei Vorliegen auffälliger Signale stimulatorisch aktiv. Erkennt das System einen bevorstehenden Anfall aufgrund der eingestellten Schwellenwerte (oder macht der Patient selbst durch Nutzung eines magnetischen Signalgebers auf einen solchen aufmerksam), werden elektrische Pulse ausgelöst. Offenbar, so zumindest die Ergebnisse einer klinischen Studie, erreicht diese Art der bedarfsgesteuerten Stimulation eine Halbierung der Anfälle und eine deutliche Lebensverbesserung aller befragten Patienten (Heck et al. 2014). Auch wenn hinsichtlich geeigneter Biomarker und passender Schwellenwerte noch viel Forschungsbedarf besteht, scheint der Weg damit vorgezeichnet für eine ganz neue Klasse autonom agierender neurotechnologischer Implantate, die unter dem Begriff Elektrozeutika (aktive Implantate) zusammengefasst werden.

Mit Blick auf den aktuellen Forschungshorizont ist zweierlei festzuhalten: Erstens, dass medizinische Anwendungsperspektiven nach wie vor den primären

⁶³ www.nebias-project.eu/ (15.12.2015)

⁶⁴ <http://sssa bioroboticsinstitute.it/news/Lifehand-bionic-hand> (15.12.2015)

Treiber der Entwicklungsdynamik bilden – dies gilt selbst für die militärische Forschungsförderung, wie die primär auf die Prothesenentwicklung fokussierten Programme der DARPA zeigen. Ein neurotechnologisch basiertes Enhancement, also eine Steigerung der Fähigkeiten gesunder Menschen, dürfte auf absehbare Zeit nur eine theoretische Option bleiben – bei invasiven Systemen verbietet es sich praktisch aufgrund der erforderlichen operativen Eingriffe, bei nichtinvasiven Systemen scheint es keinen wirklichen Zusatznutzen zu versprechen. Zweitens befinden sich selbst die fortschrittlichsten Closed-Loop-Systeme noch in einem derart frühen Entwicklungsstadium, dass laut Expertenschätzung noch mit mindestens 10 bis 20 Jahren Entwicklungsarbeit zu rechnen ist, bis an eine breite klinische Nutzung gedacht werden kann (Hofmann 2014, S.131). Dabei ist bei allen Prototypen völlig offen, ob es überhaupt gelingt, den im Fall der Neurotechnologien besonders tiefen »translationalen Graben« (translational gap) – also die Lücke zwischen Laborerfolgen und alltagstauglichen Lösungen – zu überwinden (dazu und zum Folgenden Hofmann 2014, S.125 ff.; Kübler et al. 2013; Serruya 2014). Die Geschichte der Neurotechnologien zeigt, wie groß in diesem Feld das generelle Ungleichgewicht zwischen jahrzehntelangen, intensiven Forschungsbemühungen und den eher spärlichen klinischen Anwendungserfolgen ist. Einer der Gründe ist in den relativ hohen Zulassungsanforderungen an Neurotechnologien zu sehen, die als Medizinprodukte in Deutschland unter das Medizinproduktegesetz (MPG) fallen.⁶⁵ Dieses verlangt vor Markteinführung in aller Regel den Nachweis, dass das Produkt »sicher ist, die ihm zugeschriebenen medizinischen Leistungen erbracht werden und etwaige Risiken im Zusammenhang mit der vorgesehenen Anwendung gemessen am Nutzen für den Patienten vertretbar sind« (Becker et al. 2010, S.27).⁶⁶

65 Das deutsche Medizinproduktegesetz vom 7.8.2002 setzt die europäischen Richtlinien zu aktiven Implantaten (Richtlinie 90/385/EWG), Medizinprodukten (Richtlinie 93/42/EWG) und In-vitro-Diagnostika (Richtlinie 98/79/EG) in nationales Recht um (Erstfassung 1994). Sein Zweck ist, »den Verkehr mit Medizinprodukten zu regeln und dadurch für die Sicherheit, Eignung und Leistung der Medizinprodukte sowie die Gesundheit und den erforderlichen Schutz der Patienten, Anwender und Dritter zu sorgen« (§1).

66 Soweit die Theorie, die in der Praxis aber nicht immer eingehalten wird. So ist die europäische Zulassungspraxis von Medizinprodukten in den letzten Jahren durch einige Skandale, u.a. aufgrund schadhafter Brustimplantate und Hüftprothesen, erschüttert worden und wird derzeit grundlegend revidiert. Als wesentliches Problem der aktuellen Regulierung gilt, dass für die Zertifizierung der Produkte private Prüfstellen verantwortlich sind (sogenannte benannte Stellen). Mit der neuen Medical Device Regulation (MDR), die bisher im Entwurf vorliegt, soll die Patientensicherheit durch strengere Überwachung von Herstellern und Markt gestärkt werden – eine zentrale staatliche Zulassung, wie sie für Arzneimittel erforderlich ist, ist aber nach wie vor nicht vorgesehen.



Um diese Anforderungen zu erfüllen, muss die Schnittstelle an den menschlichen Nerven verlässlich, langdauernd und funktional mit hoher Bandbreite funktionieren, bei Vollimplantaten bedeutet dies neben geringem Stromverbrauch auch verträgliche Liegezeiten im Zeitraum von Jahren sowie eine Funktionsdauer der Elektronik über Jahrzehnte. Die damit verbundenen Herausforderungen an Forschung und Entwicklung sind enorm und lassen sich nur mit einem interdisziplinären Ansatz lösen, der Materialforschung, Systemtechnik, Medizin und Neurowissenschaften einbezieht. Zu den sehr hohen Entwicklungshürden kommt verschärfend hinzu, dass der Markt für viele neurotechnologische Anwendungen sehr begrenzt ist, weil die Anwendungen zumeist nur auf sehr spezifische Indikationen zugeschnitten sind und damit auch nur für überschaubare Patientengruppen infrage kommen. So ist es nicht ausgeschlossen – und auch bereits geschehen –, dass ein Produkt zwar technologisch ausgereift und innovativ ist, aber dennoch letztendlich kommerziell scheitert – mit unter Umständen fatalen Folgen für die bereits behandelten Patienten.⁶⁷ Diese unsicheren Marktaussichten hemmen das privatwirtschaftliche Engagement und mithin auch die Innovationsdynamik.

Autonome Robotik

2.

Die Idee von Automaten, die sich menschenähnlich verhalten, lässt sich viele Jahrhunderte zurückverfolgen – in Zeiten also, in denen ihre technische Realisierbarkeit bestenfalls ansatzweise vorstellbar war, geschweige denn aufgezeigt werden konnte (Kap. II.2). Es handelte sich deshalb lange um ein Thema, das fast ausschließlich in der Kunst behandelt wurde, dabei zumeist mit stark futuristischen Anklängen (wie in der Science-Fiction). Selbst der Begriff »Roboter«, geht wie in Kapitel II erwähnt, auf ein dystopisches Drama Karel Čapeks (1920) zurück. In diesem lehnen sich die Roboter gegen ihre Ausbeutung auf – mit letztlich fatalen Konsequenzen für die Menschheit. Damit ist schon vieles angesprochen, was auch heute noch den gesellschaftlichen Diskurs um die Robotik prägt (Kap. II.3): nämlich zum einen die Auseinandersetzung mit der zunehmenden Automatisierung unserer Arbeitswelt, die bereits mit der Industrialisie-

⁶⁷ Ein Beispiel ist das implantierte Freehandsystem der Firma Neurocontrol, Cleveland (dazu und zum Folgenden Hofmann 2014, S.70). Dieses als Neuroprothese hochgelobte, durch Signale willkürlich steuerbarer Muskeln kontrollierte System ermöglichte es Tetraplegikern, funktionell sinnvolle Greifbewegungen auszuführen. Leider erreichte Neurocontrol trotz Zulassung durch die FDA, hoher Akzeptanz bei den Patienten und nachgewiesener Verbesserung bei den Tätigkeiten des täglichen Lebens keinen wirtschaftlichen Erfolg und stellte deshalb um 1998 den Verkauf ein. Eine Langzeitunterstützung der noch lebenden Nutzer ist kaum möglich und endet meist in der Explantation der Muskelelektroden und Verkabelung.



rung eingesetzt hat, zum anderen die kulturgeschichtlich tief verankerte Angst vor – und Faszination durch – menschenähnliche künstliche Wesen, wie sie beispielweise in der Legende vom Golem oder in Mary Shelleys auch durch Verfilmungen populären Roman »Frankenstein or The Modern Prometheus« zum Ausdruck kommen.

Heutzutage sind Roboter längst Realität und haben Einzug in viele gesellschaftliche Bereiche gehalten. Was genau jedoch ein Roboter ist, darüber besteht immer noch keine eindeutige Klarheit. Eine verbreitete Definition lautet, dass Roboter autonome Manipulationsmaschinen sind, die menschliche Tätigkeiten übernehmen und damit die »menschliche Handlungsfähigkeit« erweitern können (Christaller et al. 2001, S.19; RoboLaw 2014, S.15). Damit werden typischerweise folgende Merkmale in Verbindung gebracht:

- › Essenziell ist, das ergibt sich direkt aus obiger Definition, ein gewisser Grad an *Autonomie*, d.h. eine Eigenständigkeit im Verhalten, welche die Maschine befähigt, ohne externe Hilfe sinnvolle Aufgaben durchzuführen. Das setzt voraus, dass die entsprechenden Systeme über eine gewisse Wahrnehmungs- und Planungsfähigkeit verfügen.
- › Um in der Welt agieren und in diese eingreifen zu können, sollten Roboter einen realen *physischen Körper* haben, bestehend »aus mechatronischen Komponenten, Sensoren und rechnerbasierten Kontroll- und Steuerungsfunktionen« (Christaller et al. 2001, S.19). Reine Softwaresysteme oder virtuelle Maschinen sind damit keine Roboter im ursprünglichen Sinne, obwohl die Grenzen zwischen virtuellen und realen Maschinen zunehmend zu verschwimmen beginnen (etwa im Rahmen intelligenter Umgebungen).
- › Schließlich weisen Roboter in der Regel eine gewisse *Menschenähnlichkeit* auf, was auch, aber nicht unbedingt die Gestalt betrifft. Der Maschinenmensch oder humanoide Roboter dominiert zwar das populäre Bild von Robotern, wie es von der Science-Fiction-Literatur geprägt ist, ist aber technisch gesehen eher die Ausnahme. Wenn von der Menschenähnlichkeit von Robotern die Rede ist, so bezieht sich dies vor allem auf bestimmte kognitive Eigenschaften, insbesondere das Vermögen, sich intelligent, d.h. rational und zielgerichtet zu verhalten. Was das im Einzelnen bedeutet, bleibt aber selbst beim Menschen und umso mehr bei Maschinen häufig vage.

Die drei beschriebenen Kriterien – Autonomie, Verkörperung und (künstliche) Intelligenz – lassen sich, nicht zuletzt aufgrund ihrer Unbestimmtheit, einer Vielzahl von Systemen zuschreiben, die sie in höchst unterschiedlichem Maße erfüllen. Entsprechend vielfältig und weit gefächert ist das Forschungs- und Entwicklungsfeld der Robotik, das in seiner jungen, knapp 50-jährigen Geschichte verschiedenste Technologien von ganz disparater Erscheinungsform und technischen Fähigkeiten hervorgebracht hat.



Am einen Ende des Spektrums stehen die klassischen Industrieroboter, die etwa ab den 1970er Jahren im industriellen Bereich aufkamen (dazu und zum Folgenden Schaal 2014, S.3 f.). Dabei handelt es sich im Wesentlichen um große, stationäre Manipulationsarme, die einfach zu steuern sind, wenig Computerleistung brauchen und sich deshalb nur für weitgehend repetitive Tätigkeiten in hochgradig strukturierten Fabrikumgebungen eignen. Die Sensorik dieser Roboter ist minimal und bei der Konstruktion legt man Wert auf hohe Steifigkeit, Positionsgenauigkeit, Traglasten und Geschwindigkeiten. Die meisten Industrieroboter sind bis heute dieser Art. Am anderen Ende des Spektrums stehen die sogenannten autonomen Roboter, die etwa seit Mitte der 1990er Jahre erforscht und entwickelt werden. Wie der Name schon sagt, verfügen diese Systeme über einen deutlich höheren Grad an Autonomie, das heißt, sie sind fähig, wesentlich komplexere Aufgaben selbstständig durchzuführen. Das erfordert zum einen die technische Umsetzung anspruchsvoller Fähigkeiten im sensorischen, motorischen und kognitiven Bereich, eröffnet zum anderen aber auch ganz neue Anwendungsperspektiven – etwa den Einsatz im Servicebereich (Haushalt, Restaurants, Hotels etc.), in sozialen Interaktionssituationen (z. B. im Therapie- oder Pflegebereich) oder im Katastrophen- oder Kriegsfall. Bereits diese kurze Auflistung macht deutlich, dass autonome Roboter verstärkt und vor allem in enger Interaktion mit Menschen und in natürlichen Umgebungen zum Einsatz kommen. Es ist die fortschreitende Durchdringung der Gesellschaft mit zunehmend autonom agierenden Systemen dieser Art, welche die Thematik der Mensch-Maschine-Entgrenzung derzeit besonders virulent erscheinen lässt.

Es liegt auf der Hand, dass Konstruktion, Aufbau und Fertigkeiten eines autonomen Roboters spezifisch auf sein jeweiliges Einsatzgebiet zugeschnitten sein müssen und man es entsprechend auch hier mit einem sehr heterogenen Technologiefeld zu tun hat (die Bandbreite reicht vom mobilen Roboter auf vier Rädern über den Flugroboter bis hin zum winzig kleinen Nanoroboter für den Einsatz im menschlichen Körper). Als Königsdisziplin der autonomen Robotik, ja der Robotik überhaupt, gilt aber der humanoide Roboter, der den menschlichen Körper und Geist technisch nachzubilden versucht. Diese Maschinen kommen dem Menschen am nächsten, und zwar im doppelten Wortsinne – aufgrund ihrer menschenähnlichen Gestalt sind sie nämlich für Servicetätigkeiten in menschlichen Umgebungen besonders prädestiniert –, weshalb sie im Zentrum weitreichender Heils- und Schreckensvisionen stehen. Ihnen wird im folgenden Überblick deshalb auch besonderes Augenmerk geschenkt, auch und vor allem, weil sich in ihnen die fortschreitende technologische Entgrenzung von Mensch und Maschine paradigmatisch zuspitzt. Autonome Fahrsysteme und Flugroboter hingegen, die aus dieser Sicht weniger relevant erscheinen, sollen im Folgenden nicht weiter vertieft werden.

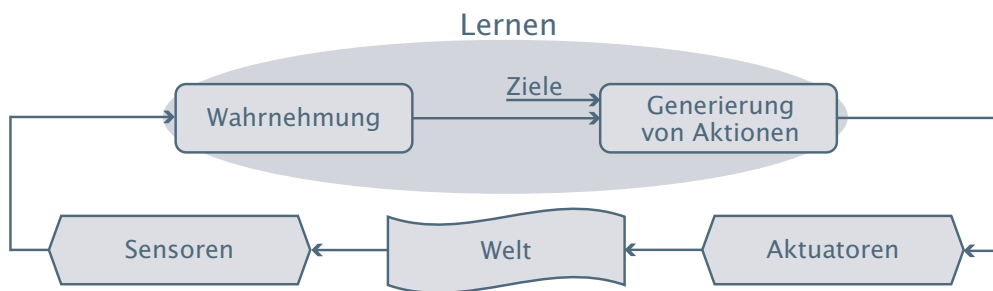
Wissenschaftliche Grundlagen: künstliche Intelligenz und maschinelles Lernen

2.1

Autonome Roboter lassen sich als Perception-Action-Learning-Systeme charakterisieren, d.h. als physische Agenten, die Wahrnehmung und Verhalten interaktiv und unter Echtzeitbedingungen realisieren (dazu und zum Folgenden Schaal 2014, S.12 f.) (Abb. III.10).

Abb. III.10

Der Perception-Action-Learning-Loop



Quelle: nach Behnke 2014, S. 10

Neben Sensoren und Aktuatoren, die es ermöglichen, die Umwelt zu vermessen und in diese einzugreifen, sind dafür insbesondere Lern- und Adaptionenmechanismen zentral, welche das System befähigen, sich an sich verändernde Umgebungen anzupassen oder komplett neue Verhaltensweisen zu erwerben. Nur so wird es auch in die Lage versetzt, robust und unfallfrei auf neue Situationen oder plötzliche Störungen reagieren zu können. Wesentliche Wissensgrundlagen für die Konstruktion derartiger Agenten sind in den beiden eng verwandten Forschungszweigen der künstlichen Intelligenz (KI) sowie des maschinellen Lernens (ML) verankert.

Künstliche Intelligenz

Das Forschungsgebiet der KI beschäftigt sich aus einer interdisziplinären Perspektive mit künstlichen Systemen, die intelligentes Verhalten zeigen, einschließlich der Fähigkeit, selbst solche Systeme zu bauen (dazu und zum Folgenden Schaal 2014, S.5 ff.). Dabei stehen nicht notgedrungen Robotersysteme oder anderweitig physische Agenten im Vordergrund, sondern ein wesentlicher, wenn nicht sogar der größere Teilbereich der KI-Forschung beschäftigt sich mit Softwaresystemen (bzw. -agenten), die Informationen rein abstrakt verarbeiten und über keinerlei mechanische Komponenten verfügen (Behnke 2014, S.5). Der KI-Forschung liegt keine feste Definition von Intelligenz zugrunde – der



Begriff ist auch beim Menschen nicht eindeutig definiert⁶⁸ –, vielmehr wird über die Beschäftigung mit künstlichen Formen der Intelligenz die Frage, was Intelligenz ist bzw. welche Aspekte intelligentes Verhalten auszeichnen, selber zum Thema (Russell/Norvig 2014).

Die KI-Forschung, deren Wurzeln bis in die 1940er Jahre zurückgehen, gilt als eine wesentliche Grundlagendisziplin für die Robotik und hat in ihrer wechselvollen Geschichte, die hier nur in Grundzügen wiedergegeben werden kann, die Entwicklung robotischer Systeme immer wieder maßgeblich geprägt. Lange war die Zielvorstellung leitend, dass es eine Intelligenz zu erschaffen gilt, die alle Aspekte der menschlichen Intelligenz repliziert: die sogenannte »starke« KI. Eine solche Maschine würde über Bewusstsein, einen freien Willen sowie Emotionen verfügen und wäre einem Menschen in ihren intellektuellen Fertigkeiten mindestens ebenbürtig. Symptomatisch für diese Sichtweise steht der von Turing im Jahr 1950 vorgeschlagene Test, mit dem festgestellt werden soll, ob eine Maschine über Denkvermögen verfügt. Die Versuchsanordnung ist folgende: Eine Person unterhält sich gleichzeitig mit einem Mensch und einem Computer, die sich in einem anderen Raum befinden und deren Antworten er folglich nicht zuordnen kann. Gelingt es dem Computer, das Gespräch so zu führen, dass der Fragesteller nicht sagen kann, ob er die Maschine ist, hat er den Test bestanden. Eine solche Maschine, so die Annahme, simuliert nicht nur Intelligenz, sie ist tatsächlich intelligent. Heutzutage hat diese umstrittene Idee der sogenannten »starken« KI innerhalb der Wissenschaft weitgehend an Anziehungskraft verloren. Stattdessen stehen spezifische Anwendungsprobleme im Vordergrund, die man mithilfe künstlicher Systeme zu lösen versucht – ohne dabei im umfassenden Sinne intelligente Maschinen schaffen zu wollen (sogenannte »schwache« KI). Die Themen der »schwachen« KI sind vornehmlich anwendungsorientiert und umfassen u. a. Wissensrepräsentation und -verarbeitung, Problemlösen und Planen, Adaption und Lernen (Russell/Norvig 2014). Nach wie vor jedoch dominiert die Vision einer »starken« KI den öffentlichen Diskurs über Robotik, beeinflusst durch die Science-Fiction-Literatur und medienwirksame Thesen einflussreicher Transhumanisten wie Ray Kurzweil (2005) (Kap. II.3).

68 Bei der Intelligenz handelt es sich um einen alltagssprachlichen Sammelbegriff, der unterschiedlichste Aspekte der kognitiven Leistungsfähigkeit des Menschen zusammenfasst. In der Psychologie wird der abstrakte Begriff der Intelligenz in der Regel mittels standardisierter Testverfahren (sogenannte IQ-Tests) messbar gemacht. Dabei gibt es eine große Zahl abweichender Operationalisierungen (sowohl für die allgemeine Intelligenz wie auch für unterschiedliche Intelligenzkomponenten), in die theoretische Hintergrundannahmen verschiedener Art einfließen. Messergebnisse wie der Intelligenzquotient, welche die intellektuelle Leistungsfähigkeit einer Person bewerten sollen, sind folglich als psychometrisches Konstrukt zu betrachten und damit von umstrittener Aussagekraft (Sauter/Gerlinger 2012, S. 53).



Die eng mit der Idee einer »starken« KI verknüpfte klassische KI-Forschung folgte mehrheitlich dem Paradigma, dass intelligentes Verhalten primär durch abstrakte Symbolverarbeitung erzeugt wird und das Substrat der Berechnungen dabei keine entscheidende Rolle spielt (Behnke 2014, S. 5). Die Entwicklung des modernen Computers, zusammen mit der Entstehung von symbolischen Programmiersprachen, die wissensbasierte Systeme unterstützen, erwiesen sich hierbei als wesentliche Triebkräfte (dazu und zum Folgenden Schaal 2014, S. 6 ff.). In den 1970er und 1980er Jahren konnte die KI-Forschung tatsächlich einige Erfolge in der abstrakten Wissensverarbeitung erzielen, wobei man sich hauptsächlich auf Gebiete konzentrierte, die logisches Problemlösen erfordern – etwa das Schachspiel oder das Beweisen von mathematischen Theoremen. In den Bereichen hingegen, in denen Menschen und selbst Kinder von Natur aus sehr kompetent sind (Wahrnehmung, Bewegung und Lernen), stießen die damaligen rechnerbasierten KI-Systeme schnell an Grenzen. Zudem war selbst das Kreieren von Expertensystemen für die logische Wissensverarbeitung so aufwendig und anwendungsspezifisch, dass die Industrie Entwicklungen im Bereich der KI weitgehend aufgab. So glänzte der Rechner »Deep Blue« von IBM, der 1997 als erster Computer gegen den damaligen Schachweltmeister Garry Kasparow gewann, vor allem dank seiner puren, rohen Rechenkraft, die durch 256 Prozessoren erzeugt wurde. Für die hochentwickelte und millionenteure Hardware gab es nach dem Wettkampf keine weitere Verwendung, sie wurde demontiert und wanderte in Teilen als Ausstellungsstück ins Museum.

Die Tatsache, dass die Fähigkeiten eines 4 Jahre alten Kindes, Objekte zu erkennen, zu manipulieren und etwas über neue Objekte zu lernen, weit über das hinausgehen, was damalige KI-Systeme erreichen konnten (und auch heute erreichen können), führte ab den 1980er Jahren zu einer zunehmenden Ernüchterung sowohl der Geldgeber als auch der Fachgemeinde (Schaal 2014, S. 6). Vor diesem Hintergrund und inspiriert von kognitionswissenschaftlichen Thesen kam zunehmend die Sichtweise auf, dass die Entwicklung von Intelligenz einen Körper braucht und sich nur in der Interaktion mit einer reichhaltigen Umwelt zeigt (»embodied cognition«) (Behnke 2014, S. 5; Pfeifer/Bongard 2007). Diese Embodimenthypothese leitete in zweierlei Hinsicht einen Paradigmenwechsel in der KI-Forschung ein (dazu und zum Folgenden Schaal 2014, S. 7 f.):

- › Zum einen rückte die Robotik (im Speziellen die humanoide Robotik), die sich mit physisch (und nicht nur formal) repräsentierten Systemen befasst, wieder verstärkt in den Fokus der KI-Forscher. Zwar hatte die Robotik schon zuvor eine wichtige Rolle in der KI gespielt, jedoch meist in Form mobiler Roboter, die sich auf Rädern fortbewegen und über keine Manipulatoren verfügen. Die Steuerung von mobilen Robotern ist jedoch verhältnismäßig trivial, da sie auf den zweidimensionalen Raum beschränkt ist, sodass diese sich besonders zur Untersuchung von einfachen Wahrnehmungs- und Planungs-



problemen anbieten. Als Modell zur Erforschung von komplexeren (menschlichen) Aspekten der Intelligenz sind mobile Roboter jedoch weniger geeignet. Der wesentlich bessere Ansatzpunkt hierfür sind humanoide Roboter, die in Echtzeit die nichttriviale Aufgabe des Balancierens in hochdimensionalen Räumen lösen und komplexe Bewegungssequenzen planen müssen. Diese komplexen Anforderungen benötigen oft andere Algorithmen als diejenigen, die in der KI der mobilen Robotik verwendet werden.

- › Der neue Blick auf die situative Interaktion eines verkörperten Agenten mit seiner Umwelt hat zum anderen dazu beigetragen, dass neben der abstrakten Intelligenz andere Verhaltensaspekte stärker in den Fokus der KI-Forschung geraten sind. Stand im klassischen Paradigma noch der abstrakt »denkende« Rechner im Vordergrund, ist es nun der »lernende« Akteur, der sich mit seinem Körper in Echtzeit in einer neuen Umwelt zurechtfinden muss. Gleichzeitig orientierte man sich am Vorbild realer Organismen, speziell der neuronalen Informationsverarbeitung, um solche Lernprozesse künstlich zu simulieren (Görz et al. 2014, S. 12 f.). Die Forschung an »künstlichen neuronalen Netzen« geht bis in die 1940er Jahre zurück, erlebte jedoch in den 1980er Jahren einen anhaltenden Aufschwung und begründete letztlich die Entwicklung einer neuen Teildisziplin der KI-Forschung, dem ML, welche die Robotikforschung der letzten Jahre tiefgreifend bestimmt hat.

Maschinelles Lernen

Die Wurzeln des ML liegen in der KI-Forschung, wie zuvor beschrieben, der Bereich hat sich jedoch längst als eigenständiges Forschungsfeld etabliert und von seinen disziplinären Ursprüngen emanzipiert. Sehr allgemein gesprochen wird sich beim ML mit dem Generieren und dem Studium von Algorithmen, die sich datengetrieben verbessern können, befasst (Schaal 2014, S.9). Der Schwerpunkt liegt somit auf empirischen, datengetriebenen Methoden und insbesondere der statistischen Datenverarbeitung und nicht, wie bei der herkömmlichen KI-Forschung, auf der logischen Wissensverarbeitung. Wesentliche Grundlagen liegen in der Theorie der »künstlichen neuronalen Netze«, die sich – wie der Name bereits besagt – von Strukturmerkmalen des Gehirns inspirieren ließ, um Lernvorgänge mathematisch zu modellieren. Die Grundidee ist, durch die Verknüpfung einfacher informationsverarbeitender Einheiten (den künstlichen Neuronen) assoziatives Lernen zu ermöglichen, indem die Gewichtung der Verknüpfungen und damit der Output des Netzes sich erfahrungsabhängig verändert (Behnke 2014, S.5; Oberhofer/Zimmerer 1996).

Indem es gelang, Netzwerke mit vielen Schichten zu generieren, die theoretisch unbegrenzte Lernfähigkeit besitzen und die einfach zu trainieren sind (Bishop 2007), kam es Mitte der 1980er Jahre zu einer regelrechten Explosion



der Forschungsaktivitäten (dazu und zum Folgenden Schaal 2014, S.9 f.). Plötzlich schienen mithilfe neuronaler Netze Forschungsfragen lösbar, an denen die KI-Forschung bislang gescheitert war, wie zum Beispiel Probleme beim maschinellen Sehen (z. B. Objekterkennung bzw. Segmentierung), bei kognitiven Prozessen (z. B. Gedächtnis und Assoziationen) oder der Sprachverarbeitung etc. (Rumelhart/McClelland 1986). Die ersten empirischen Erfolge weckten gleichzeitig das Interesse von stärker mathematisch und statistisch orientierten Forschern, die Anfang der 1990er Jahre aufzeigten, dass neuronale Netze als statistische Algorithmen interpretiert werden können (Hertz et al. 1991). Damit verlor aber auch die »neuronale Metapher«, d.h. die biologische Interpretation der abstrakten Netzwerke aus künstlichen Neuronen, in der Forschung immer mehr an Bedeutung (Suzuki 2013). So kam es, dass das Forschungsgebiet der neuronalen Netze erst dem »statistischen Lernen« und schließlich allgemein dem ML zugeordnet wurde.

Heute gehören das ML im Allgemeinen und künstliche neuronale Netze im Speziellen zu den wichtigsten Werkzeugen der Informationstechnologie. Ein zentrales Anwendungsfeld ist die Mustererkennung, wie sie bei der Datenverarbeitung und -klassifikation (etwa Bildklassifikation, Sprachverarbeitung), in der Bioinformatik (z. B. bei genetischen Analysen) oder in den Neurowissenschaften (etwa der EEG-Signalverarbeitung) eine zentrale Rolle spielt – die potenziellen Anwendungsmöglichkeiten sind nahezu unbegrenzt. Insbesondere zwei Entwicklungen haben dazu beigetragen, dass sich Methoden des ML in immer mehr Bereiche auszubreiten beginnen (dazu und zum Folgenden Schaal 2014, S. 10):

- › Durch die rasante Entwicklung der Computer- und Speicherleistung bietet sich erstens die Möglichkeit, immer umfangreichere Netzwerkarchitekturen zu entwickeln (bzw. die gleiche Rechenkraft lässt sich entsprechend kostengünstiger realisieren).
- › Zweitens liegt mit den enormen Datenmengen, die über das Internet zugänglich sind, ein potenziell riesiger Schatz für statistische Analysen bereit (Big Data), den vor allem Großkonzerne wie Google, Facebook, Amazon, Apple etc. durch Data-Mining zu erschließen suchen und für die Weiterentwicklung ihrer ML-Algorithmen nutzen.

Das Training mehrschichtiger Architekturen an großen Datenmengen wird auch als »Deep Learning« bezeichnet – ein Begriff, den Hinton/Salakhutdinov (2006) geprägt haben (dazu und zum Folgenden Behnke 2014, S.30). Häufig werden hierbei zunächst durch unüberwachtes Lernen Merkmalshierarchien erzeugt, die dann durch diskriminatives Training an die konkrete Aufgabe angepasst werden – die Berechnungen sind hochgradig parallel und können gut durch parallele Hardwarearchitekturen beschleunigt werden. Auf diese Weise konnten in den letzten Jahren einige teils spektakuläre Erfolge erzielt werden.



So basiert Apples Spracherkennungssoftware »Siri« auf Deep-Learning-Algorithmen. Im Sommer 2014 konnte Google mit einer besonders tiefen Netzwerkarchitektur die Erkennungsleistung bei zwei ImageNet-Challenges – Wettbewerben zur Bildkategorisierung und zur Objektdetektion in Bildern – deutlich verbessern (Szegedy et al. 2014). Kürzlich wurde zudem, ebenfalls von einem Google-Team, ein Verfahren zur natürlichsprachlichen Beschreibung von Bildern vorgestellt, das erstaunlich oft plausible Bildbeschreibungen produziert (Vinyals et al. 2014). Wenn es noch eines Beweises für das große Potenzial des ML bedurft hätte, dann wurde er im März 2016 endgültig geliefert, als das Programm AlphaGo der Google-Firma DeepMind einen der weltbesten Go-Profispieler in einem Wettkampf vernichtend schlug. Dieser Erfolg ist deshalb so bemerkenswert, weil Go wesentlich komplexer und weniger berechenbar ist als Schach und deshalb noch auf längere Zeit als eine der letzten Bastionen menschlicher Überlegenheit gegolten hatte.⁶⁹ Dass diese Bastion nun genommen wurde, und zwar auf eine auch für Experten durchaus überraschende und beeindruckende Weise, fand großes mediales Echo.

Klassische Lernaufgaben in ML zeichnen sich in der Regel dadurch aus, dass das System erst an einem möglichst umfangreichen Datensatz trainiert werden muss, bevor der Lernerfolg anschließend an Testdatensätzen überprüft werden kann (Schaal 2014, S. 10 f.). Viele Methoden sind darauf angewiesen, dass ein Experte Daten auswählt, geeignete Vorverarbeitungsschritte zur Merkmalsextraktion implementiert, den Lernalgorithmus und dessen Parameter bestimmt und die für die Informationsverarbeitung genutzten Repräsentationen des Systems festlegt (Behnke 2014, S. 29). Mit anderen Worten, die Autonomie der Lernsysteme ist bislang sehr begrenzt und die bisherigen Erfolge beschränken sich auch weitgehend auf Bereiche, in denen vorstrukturierte Datenmengen vorliegen (Toussaint et al. 2010, S. 3). Bei der Robotik, besonders der autonomen Robotik, sind diese Rahmenbedingungen in der Regel nicht oder nur unzureichend gegeben, denn ein Roboter, der mit einer komplexen Umwelt interagiert, hat eben nicht allein mit vorgefertigten Daten zu tun, sondern es ist geradezu erwünscht, dass er selbstständig aus eigener Erfahrung lernt. Eine wesentliche Aufgabe von ML in der Robotik besteht daher darin,

⁶⁹ AlphaGo kombiniert verschiedene KI-Lernverfahren: zum einen das auch bei herkömmlichen Go-Programmen übliche Monte-Carlo-Verfahren, das verschiedene Spielverläufe statistisch simuliert; zum anderen tiefe neuronale Netzwerke, welche die Stellung bewerten und die Suche nach dem nächsten Zug eingrenzen (Silver et al. 2016). Die neuronalen Netze wurden erst anhand von Meisterpartien trainiert und lernten anschließend weiter dazu, indem sie tausende Partien gegen sich selbst spielten. Beeindruckend ist, in welcher kurzen Zeit das Programm auf diese Weise, also quasi selbstlernend, seine Spielstärke vom Amateur- auf das Profiniveau heben konnte. Die Hoffnung der Entwickler ist, dass die zugrundeliegende Lernarchitektur keineswegs nur auf Go beschränkt ist, sondern sich auch auf gesellschaftliche Probleme anwenden lässt.



Lernverfahren zu entwickeln, die ein erstrebtes Zielverhalten aus der Interaktion mit der Umwelt erzeugen können. Besonders relevant sind in diesem Zusammenhang Methoden des Lernens aus Belohnungen und Bestrafungen, das sogenannte Reinforcementlernen (Behnke 2014, S. 13; vgl. Sutton/Barto 1998). Diese Methoden ändern das Roboterverhalten so, dass der erwartete Nutzen maximiert wird. Es muss also nicht ein korrektes Verhalten für jede Situation vorgegeben werden, sondern die Lernziele werden als Kosten- oder Belohnungsfunktion definiert – beispielsweise zur Optimierung von Gangparametern eines zweibeinigen Roboters (Faber/Behnke 2007). In der Praxis stößt dieses Lernparadigma jedoch auf enorme Herausforderungen, die sich im sogenannten »Exploration-Exploitation-Dilemma« (Thrun 1992) zuspitzen (dazu und zum Folgenden Schaal 2014, S. 11 f.). Der Roboter lernt nur, wenn er Neues ausprobiert, was allerdings die Gefahr birgt, dass durch suboptimales Verhalten Schäden am Roboter oder dessen Umgebung entstehen. Somit ergibt sich ein Zielkonflikt zwischen dem Ausprobieren von neuem Verhalten (Exploration) einerseits sowie der Nutzung von bereits erlerntem Verhalten (Exploitation) andererseits. Die sich daraus ergebenden Fragestellungen, die sich darin zuspitzen, dass Lerndaten selbst erzeugt werden müssen, gehen über das Fachgebiet des ML hinaus, in dem normalerweise mit fixen Trainingsdaten gearbeitet wird, und benötigen daher eigene Algorithmen.

Auffällig ist, dass spätestens mit dem jüngsten Aufkommen von Deep-Learning-Algorithmen, deren Struktur der Hierarchie visueller Areale im Kortex ähnelt (Behnke 2014, S. 30), biologische Analogien wieder stärker in den Vordergrund rücken. Nach wie vor interessieren sich Robotik- und KI-Forscher für biologische Vorbilder und speziell die Architektur des Gehirns, um informationstechnologische Systeme weiter zu verbessern (»reverse engineering of the brain«). Diese haben jedoch inzwischen einen Entwicklungsgrad erreicht, der umgekehrt auch verstärkt auf neurowissenschaftliches Interesse stößt (Rucci et al. 2007). So werden Roboter zunehmend als experimentelle Plattform benutzt, um Aspekte der menschlichen Informationsverarbeitung, kognitive Prozesse oder Sprachentwicklung modellhaft untersuchen zu können. Hierfür sind vor allem humanoide Roboter geeignet, welche über komplexe sensomotorische Fähigkeiten verfügen, die denen des Menschen ähnlich sind – Beispiele für entsprechende Forschungsplattformen sind der Roboter »iCub« und der in Berlin entwickelte Roboter »Myon« (Kasten) (Hild et al. 2012; Metta et al. 2008). Dementsprechend ist aktuell eine zunehmende Konvergenz zwischen neurowissenschaftlichen, kognitionswissenschaftlichen und informationstechnologischen Fragestellungen zu beobachten, die auch im Aufblühen hybrider Forschungsbereiche wie der kognitiven Robotik, der Neuroinformatik oder der Neurorobotik zum Ausdruck kommt. Diese Entwick-



lung ist bislang hauptsächlich auf die Grundlagenforschung beschränkt und wird etwa im aktuellen Human Brain Project der EU stark vorangetrieben.

Myon: Ein Roboter als Opernstar?

Myon ist ein humanoider Roboter, der von einer Forschergruppe um Prof. Manfred Hild am Forschungslabor Neurorobotik der Beuth Hochschule für Technik Berlin konzipiert und realisiert worden ist. Der 1,25 m große Roboter verfügt über ein innovatives Design: Er ist modular aufgebaut und kann während des Betriebs auseinandergebaut werden, da die einzelnen Module (Arme, Beine, Kopf und Torso) über eine eigene Energieversorgung und Steuerungseinheit verfügen und damit völlig unabhängig voneinander funktionieren. Myon besteht insgesamt aus 48 Aktuatoren, die 32 Freiheitsgrade bereitstellen. Die Welt nimmt er u. a. über eine am Kopf angebrachte Kamera wahr, anstelle der Ohren befinden sich Mikrofone.

Das menschenähnliche Äußere hat System, denn Myon ist in erster Linie ein Forschungsobjekt, anhand dessen Fragen aus dem Bereich der sogenannten Neurorobotik untersucht werden sollen. Dieses neue Forschungsfeld, das Methoden aus den Neurowissenschaften, der Robotik und der KI-Forschung kombiniert, interessiert sich dafür, wie autonome Roboter (auf Basis neuronaler Regelungsstrukturen) ihr Verhalten adaptiv an die reale Umwelt anpassen. Auch das Innenleben von Myon wurde deshalb in Analogie zur menschlichen Biologie gestaltet, indem »die sensomotorischen Regelkreise ... in einer gewissen Abstraktionsebene die Strukturen und Neurodynamiken eines Gehirns nachbilden«.70 Nach Aussage von Prof. Manfred Hild befindet sich Myon derzeit noch etwa auf dem Entwicklungsstand eines Kleinkindes – die Hoffnung ist, dass er mit der Zeit selbstständig neue Fähigkeiten hinzugewinnt und sich schrittweise zu einem intelligenteren Wesen weiterentwickelt.71

Voraussetzung dafür ist, dass er fortlaufend neue Erfahrungen macht und neuen Eindrücken ausgesetzt wird. Ein Experiment, das in diesem Sinne durchgeführt wurde, machte Myon auch einem breiteren Publikum bekannt. In der Oper »My Square Lady«, die in der Komischen Oper Berlin im Juni und Juli 2015 aufgeführt wurde, spielte Myon die Hauptrolle. Dadurch soll er »das ›Kraftwerk der Gefühle‹ Oper in all seinen Facetten erkunden und dabei lernen, was es heißt, menschliche Gefühle zu empfinden, sie auszudrücken und bei anderen hervorzurufen«.72

70 www.neurorobotics.de/index_de.php (15.12.2015)

71 www.berliner-zeitung.de/wissen/roboter-myon-lernt-wie-ein-kind,10808894,25535008.html (15.12.2015)

72 www.komische-oper-berlin.de/spielplan/my-square-lady/ (15.12.2015)



Nach zweijährigen Proben, die als offener Lernprozess angelegt waren, erfolgte dann der große Bühnenauftritt – mit großem Presseecho, aber insgesamt wohl eher bescheidenem Ergebnis, wie ein Videomitschnitt der Aufführung zeigt.⁷³

Wo steht die autonome Robotik?

2.2

Ein autonomer Roboter ist eine hochkomplexe Maschine, die aus einer Vielzahl an Komponenten besteht, wozu Sensoren zur Wahrnehmung der Umwelt, Steuerungsmechanismen zur Realisierung der Aktionsmöglichkeiten sowie Lern- und Planverfahren gehören. Die Entwicklung solcher Systeme benötigt eine konsequent interdisziplinäre Herangehensweise, die weit über die skizzierten Grundlagenfächer der KI- und ML-Forschung hinausgeht und vor allem auch technische Disziplinen wie die Mechatronik, die Regelungstechnik oder die Elektrotechnik involviert. Eine besondere Herausforderung stellt die Integration der unterschiedlichen Systemkomponenten zu einem funktionierenden Gesamtsystem dar. Der aktuelle Stand von Forschung und Entwicklung in den Bereichen Wahrnehmung, Lern- und Planverfahren (Steuerung) sowie Systemintegration soll im Folgenden genauer beleuchtet werden.

Wahrnehmung

Grundlage für die Verhaltenssteuerung autonomer Roboter ist die Erstellung eines hinreichend guten Modells der Roboterumgebung und des eigenen Zustands aus sensorischen Informationen (Behnke 2014, S.6). Dafür braucht es zweierlei: erstens vielfältige Sensoren, die bestimmte Eigenschaften des Roboters und der Umgebung erfassen können, und zweitens Verfahren, die eine semantische Interpretation der wahrgenommenen Objekte ermöglichen (Objektklassifikation).

Moderne Robotersysteme enthalten unterschiedlichste Wahrnehmungssensoren (dazu und zum Folgenden Behnke 2014, S.6 f.). Der Wahrnehmung der Postur – der relativen Lage der einzelnen Teile des Roboters zueinander – dienen Gelenkencoder, die Position und Geschwindigkeit messen. Für die Schätzung der Lage im Raum werden Beschleunigungs-, Drehraten- und Magnetfeldsensoren kombiniert (Allgeuer/Behnke 2014). Zur Messung von Interaktionskräften werden an Berührungspunkten zwischen Roboter und Umwelt Kraftmomentensensoren verbaut. Künstliche Häute ermöglichen die Erfassung hap-

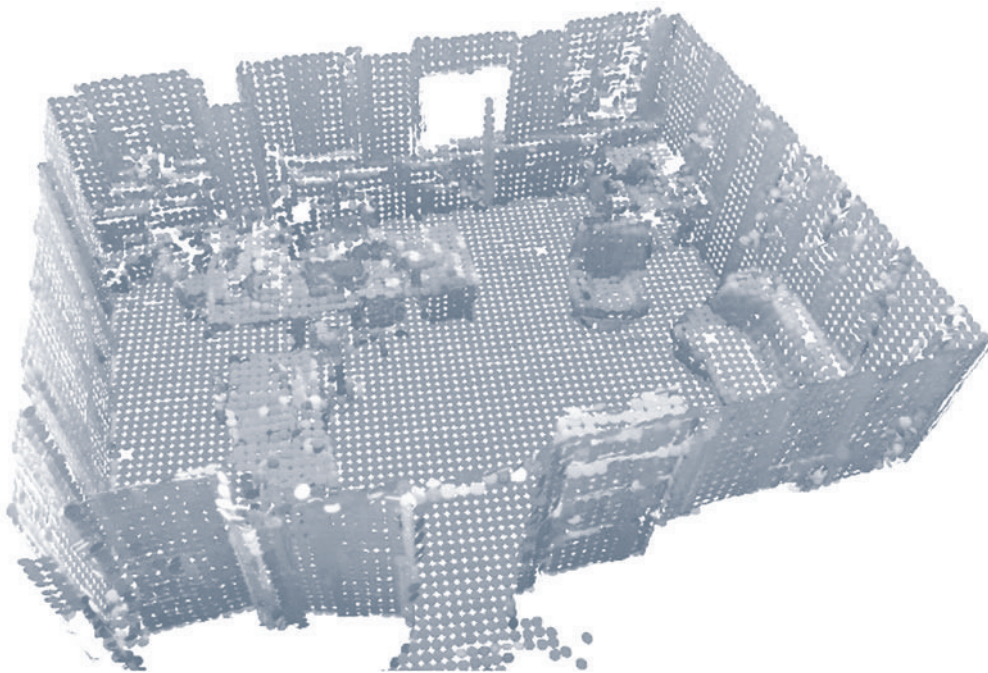
⁷³ www.youtube.com/watch?v=G13lHqIjoRA (15.12.2015)



tischer Reize. Mikrofone dienen der akustischen Erfassung von Szenen oder der Spracherkennung (Bregman 1994; Rabiner/Juang 1993). Autonome Roboter, die sich möglichst selbstständig bewegen sollen, brauchen alle diese sensorischen Möglichkeiten, zweifelsohne am wichtigsten jedoch sind visuelle Wahrnehmungsmöglichkeiten, die ein möglichst detailliertes Abbild der Umgebung erzeugen. Im Falle mobiler Roboter, die sich auf Rädern fortbewegen, ist die einfache Erfassung der horizontalen Ebene häufig ausreichend. Dazu dienen beispielsweise Sensorsysteme zur Abstandsmessung wie Ultraschallsensoren oder 2-D-Laserscanner, mit welchen die Distanz zu Hindernissen bestimmt werden kann. Sobald aber im Raum operiert wird (was bei den meisten autonomen Robotern der Fall ist), ist eine angemessene räumliche Wahrnehmung unerlässlich, damit jedem Objekt (inklusive des Roboters selbst) auch eine Position und eine Orientierung im Raum zugeordnet werden können – nur so lassen sich Handlungen in Bezug auf diese Objekte planen. Zum Einsatz kommen hierbei beispielsweise schwenkbare Kamerasysteme oder Laserscanner, die Abstände in einem weiten Sichtfeld – oder gar in alle Raumrichtungen – messen können. Abbildung III.11 zeigt eine mit einem schwenkbaren Laserscanner erzeugte Oberflächenelementkarte eines Innenraums.

Abb. III.11

Mit schwenkbarem Laserscanner erzeugte
3-D-Karte eines Innenraums



Quelle: Kläß et al. 2012



Um zielgerichtet navigieren zu können, muss ein Roboter in der Lage sein, sowohl die verschiedenen Messungen zu Karten zu aggregieren als auch gleichzeitig seine eigene Position und Orientierung innerhalb dieser Karten zu lokalisieren (dazu und zum Folgenden Behnke 2014, S.8, sowie Schaal 2014, S.32 f.). Diese zirkuläre Aufgabe, die auch als »Simultaneous Localization and Mapping« (SLAM) bekannt ist, wird dadurch erschwert, dass sowohl die sensorischen Messungen rauschbehaftet und unzuverlässig sind, als auch die Roboterbewegung (und nicht allzu selten auch die Bewegung der Objekte selber) die Unsicherheit erhöht. Unerlässlich ist vor allem eine ausreichende Tiefenwahrnehmung. Ein übliches Verfahren, das sich am menschlichen Sehsystem orientiert, ist die gleichzeitige Aufnahme der Szene aus mehreren Perspektiven, etwa durch Stereokamera-paare. Die Stereowahrnehmung ist allerdings ziemlich rechenaufwendig und nicht sehr zuverlässig. Ganz neue Möglichkeiten bieten dagegen RGB-D-Kameras, die infrarote Lichtprojektionen auf die Umgebung werfen und so texturierte Oberflächen für die Tiefentriangulation schaffen. So liefert der Kinectsensor der Firma Microsoft, der für Videospiele entwickelt wurde und schon für rund 200 Euro erhältlich ist, direkt 3-D-Bilddaten mit recht hoher Qualität. Für die 3-D-Wahrnehmung in der Robotik hat dies in den letzten 5 Jahren einen breiten Umschwung eingeleitet, da es viel einfacher geworden ist, Manipulationsobjekte in der Umwelt zu erkennen und zu lokalisieren. Nahezu jedes Robotiklabor benutzt heute diese Sensoren. Im Außenbereich sind RGB-D-Kameras aufgrund der Überstrahlung mit Sonnenlicht jedoch nicht einsetzbar.

Um ein robustes Abbild der Außenwelt zu erhalten, ist zusätzlich zu der 3-D-Rekonstruktion in der Regel auch eine semantische Interpretation der Umgebung erforderlich, das heißt die Unterscheidung und Kategorisierung der wahrgenommenen Objekte (Behnke 2014, S.8). Die zuvor beschriebenen Verfahren, die dem Deep Learning zuzuordnen sind, erzielen hierbei bereits sehr gute Ergebnisse. Voraussetzung dafür sind allerdings ein umfangreiches Training anhand von Lerndaten und der Zugriff auf große Rechen-netzwerke, was sich bei Robotern nicht so einfach umsetzen lässt. Die autonome Wahrnehmung in Echtzeit, vor allem in Bezug auf ein umfassendes Szenenverständnis auf Grundlage einzelner Wahrnehmungsinhalte, stellt immer noch eine große Herausforderung dar (Toussaint et al. 2010, S.14). Dies gilt im Besonderen für die »Action Recognition«, das heißt das Erkennen von Handlungen (dazu und zum Folgenden Schaal 2014, S.32). Speziell geht es darum, menschliche Bewegungen zu erkennen und dann die Absicht dieser Bewegungen zu eruieren. So ist es relativ einfach, eine Fortbewegung auf zwei Beinen zu erkennen. Schon wesentlich schwieriger zu erkennen ist, ob jemand Fußball spielt oder joggt. Letztendlich möchte man erreichen, dass autonome Roboter die Fähigkeit bekommen, jede Art von Handlung zu verstehen, wie zum Beispiel das Greifen einer Tasse, das Essen mit einer Gabel oder das Spielen mit Bauklötzen. Das ist



besonders dann essenziell, wenn autonome Roboter in menschlichen Umgebungen und mit Menschen interagieren sollen. Um diese äußerst anspruchsvolle Fähigkeit realisieren zu können, wird eine Kombination aus Situationsverständnis, Objekterkennung, dem Erkennen von Handlungsabsichten und dem Verstehen von zeitlichen Abläufen benötigt.

Auch wenn die Sensorik und die Methoden der Dateninterpretation, Mustererkennung und Kartierung in den letzten Jahren deutliche Fortschritte erzielt haben, ist abschließend festzuhalten, dass die Wahrnehmung der Umwelt immer noch ein großes Problem für autonome Roboter darstellt (dazu und zum Folgenden Behnke 2014, S.9). Schon die sensorische Erfassung ist nicht immer einfach. Transparente Materialien, reflektierende Oberflächen, homogene Flächen, tief schwarze Oberflächen sowie kleine bzw. entfernte Objekte sind schwer zu erfassen. Häufig müssen verschiedene Sensormodalitäten eingesetzt werden, um Defizite einzelner Modalitäten auszugleichen, was wiederum eine Herausforderung an die Datenfusion darstellt. Auch die Interpretation der Messungen ist noch immer nicht gut genug gelöst. Dies liegt vor allem an der Variabilität der vorkommenden Szenen. Unterschiedliche Ausprägungen von Objektkategorien, artikulierte und deformierbare Objekte, Änderungen in Perspektive und Beleuchtung, Verdeckungen etc. machen die Interpretation schwierig. Menschen können vielfältige vage Hinweise, z. B. zur Rekonstruktion der 3-D-Geometrie, integrieren und sind der maschinellen Wahrnehmung immer noch in vielen Aspekten voraus. Nur wenn die Szenen relativ einfach sind und die Aufnahmebedingungen kontrolliert werden können, wie beispielsweise bei der Erkennung von Fahrzeugkennzeichen, der Postsortierung oder der Qualitätskontrolle in der Industrie, sind heutige Wahrnehmungssysteme dem Menschen ebenbürtig oder sogar überlegen.

Lern- und Planverfahren

Roboter verfügen üblicherweise über eine Vielzahl von Aktionsmöglichkeiten (dazu und zum Folgenden Behnke 2014, S.9 ff.). Sie können sich in ihrer Umgebung bewegen, Objekte manipulieren und akustische oder optische Signale von sich geben. Vor diesem Hintergrund ist die Aufgabe der Verhaltensplanung, auf Basis der durch die Wahrnehmung gewonnenen Weltsicht konkrete Aktionen zu generieren, welche die Welt – inklusive den Roboter selbst – so beeinflussen, dass bestimmte Ziele erreicht werden. Klar ist dabei, dass autonome Roboter nicht vollständig für alle Anforderungen unterschiedlicher Einsatzumgebungen vorprogrammiert werden können. Daher ist in diesem Zusammenhang auch die Implementierung von Lernverfahren für die Anpassung an konkrete Gegebenheiten erforderlich.



Bei der Verhaltenssteuerung lassen sich drei grundsätzliche Herangehensweisen unterscheiden, nämlich deliberative, reaktive sowie hybride Kontrollstrukturen (dazu und zum Folgenden Behnke 2014, S. 10 f.):

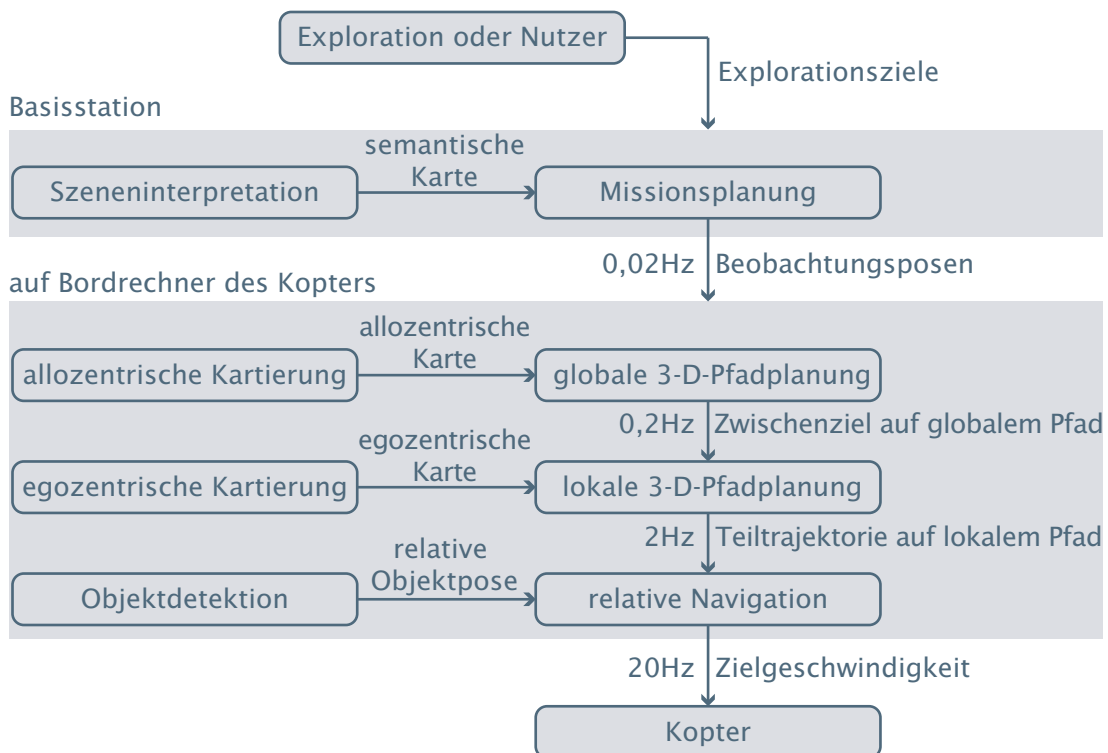
- › Roboter mit sogenannter deliberativer Kontrollarchitektur versuchen erst, ein möglichst vollständiges Modell des für die Anwendung relevanten Teils der Welt zu erstellen. Darauf basierend wird eine Aktionsfolge geplant, die einen Startzustand in einen Zielzustand überführt, bevor es schließlich zur Ausführung der Aktionen kommt. Dieses stark planungsbasierte, von der klassischen KI-Forschung geprägte Vorgehen benötigt viel Rechenzeit und gute Modelle der Welt und der Effekte von Aktionen. Letzteres ist jedoch schwierig zu erreichen, da die Welt in der Regel nicht deterministisch aufgebaut ist: Messungen sind rauschbehaftet und können fehlerhaft sein, Aktionen führen nicht sicher zu Nachfolgezuständen, sondern Nachfolgezustände werden nur mit bestimmten Wahrscheinlichkeiten erreicht.
- › Reaktive Ansätze beruhen auf einem einfachen Stimulus-Response-Modell und sind dadurch weniger planungs-, dafür stärker verhaltensbasiert. Konkret heißt dies, dass der Roboter direkt auf sensorische Inputs reagiert, ohne dass eine übergeordnete Planung erzeugt wird – generiert wird also eine augenblickliche Abbildung, die zukünftige Entwicklungen nicht berücksichtigt. Gegenüber deliberativen Steuerungsansätzen hat dies den Vorteil, dass so schnell auf neue Situationen reagiert werden kann, allerdings ist die Komplexität der erzeugbaren Verhaltensweisen begrenzt.
- › Bei hybriden Ansätzen wird versucht, das Beste aus reaktiven und deliberativen Zugängen zu vereinen, meist basierend auf einer hierarchischen Vorgehensweise. Ein typisches Beispiel hierfür ist eine dreischichtige Architektur (Gat 1998). Diese besteht aus einer langsamen deliberativen Ebene, die in die Zukunft plant, einer schnellen reaktiven Ebene, die primitive Verhaltensweisen generiert, und einer Zwischenschicht, die zwischen beiden vermittelt. Wie in Abbildung III.12 illustriert, werden in aktuellen Systemen auch mehr als drei Schichten benutzt. So plant der von der Forschergruppe »Mapping on Demand« entwickelte Flugroboter anhand von Explorationszielen Missionen und konkretisiert diese schrittweise (Nieuwenhuisen/Behnke 2014). Wichtig ist dabei, dass die niedrigeren Ebenen immer aktuellere und detailliertere Repräsentationen benutzen sowie Ausnahmesituationen erkennen und diese der nächsthöheren Ebene signalisieren.

Auch wenn für die Generierung des Roboterhaltens schon brauchbare Ansätze existieren, müssen oft vereinfachende Annahmen gemacht werden, um das Roboterhalten in Echtzeit planen zu können – beispielsweise, dass die Dynamik der Roboterbewegung keine wesentliche Rolle spielt (Behnke 2014, S. 12). Bei komplexen Bewegungsabläufen, wie sie humanoide Manipulationsro-



boter zu vollbringen haben, stoßen die derzeitigen Verfahren definitiv an ihre Grenzen. Derartige Roboter verfügen oft über mehr als 50 Regelungsdimensionen, darunter sieben Freiheitsgrade pro Arm und Bein und darüber hinaus zehn bis 20 zusätzliche Freiheitsgrade für Körper, Finger und Kopf (dazu und zum Folgenden Schaal 2014, S.29 f.). Um komplexes Verhalten zu verwirklichen – etwa Manipulationsaufgaben, Fortbewegung auf Beinen oder Ganzkörpermanipulationen (Aufgaben, bei denen Arme und Beine gleichzeitig im Einsatz sind, wie z. B. beim Aufheben eines Objekts vom Boden) –, ist deshalb noch viel manuelle Programmierung und externe Feinsteuerung vonnöten. Damit beispielsweise der Roboter eine Tasse aus dem Küchenschrank holt, wird die Gesamtaufgabe in einzelne Unteraufgaben unterteilt (z. B. »in die Küche gehen«, »den Schrank finden«, »die Schranktür öffnen«, »die Tasse greifen« etc.) und die Abfolge definiert, in denen sie durchgeführt werden sollen. Anschließend wird das Verhalten des Roboters für die einzelnen Unteraufgaben programmiert oder idealerweise durch Lernverfahren erzeugt.

Abb. III.12 Geschichtete Kontrollarchitektur eines Flugroboters



Quelle: nach Nieuwenhuisen/Behnke 2014

Letztere sind wichtig, um Anpassungen an konkrete Gegebenheiten zu ermöglichen und zukünftiges Verhalten möglichst zu verbessern (dazu und zum Folgen-



den Behnke 2014, S.12 u. 14, sowie Schaal 2014, S.11 f.). Wie bereits erwähnt, greift man häufig auf Kostenfunktionen zurück, um Lernziele zu definieren (sogenanntes Reinforcementlernen).⁷⁴ Aufgrund der Komplexität des Gesamtsystems, das unterschiedliche Komponenten und damit Lernaspekte beinhaltet, geschieht dies häufig nicht auf der Ebene des Gesamtsystems, sondern unterschiedliche Teile des Systems – von der Wahrnehmung bis zur Aktionsgenerierung – benötigen spezielle Lernalgorithmen. Ob das gewünschte Lernergebnis dann jeweils auch erreicht wird oder gar ein komplettes System mit Wahrnehmungs- und Handlungsfähigkeit entwickelt werden kann, steht und fällt zum einen mit der Qualität der Kostenfunktionen – diese passend zu formulieren, ist Aufgabe von Experten und alles andere als trivial. Zum anderen sind die Verfahren des Reinforcementlernens gerade in komplexen Umwelten sehr zeitaufwendig, sodass eine große Zahl an Erfahrungen (bzw. Misserfolgen) benötigt wird, um das optimale Ergebnis zu erreichen. Aus all diesen Gründen ist es bis zur Realisierung von praxistauglichen Robotern, die selbstständig aus eigener Erfahrung lernen, noch ein weiter Weg. Derzeit werden Lernverfahren in der Regel nur zur Optimierung einzelner Systemkomponenten eingesetzt, bevor der Roboter zur Anwendung kommt. Ein Weiterlernen beim Anwender im täglichen Einsatz – ohne Überwachung durch den Hersteller – ist mit der Gefahr verbunden, dass sich das Roboterverhalten unvorhersehbar ändert, und wird daher vermieden.

Autonom agierende Roboterhelfer, wie sie zuvor als Vision beschrieben wurden, sind folglich noch weitgehend Zukunftsmusik (dazu und zum Folgenden Schaal 2014, S.30, sowie Behnke 2014, S.30 f.). Gewisse autonome Unterverhalten sind zwar heute schon erzielbar, wie z. B. das Balancieren und Gehen auf zwei Beinen im ebenen Gelände oder das Greifen von Gegenständen. Bei komplexeren Szenarien stößt die Steuerung autonomer Roboter jedoch noch schnell an Grenzen. So ist in vielen Fällen überhaupt nicht klar, wie das Ziel eines Verhaltens formalisiert werden kann – was bedeutet es beispielsweise mathematisch, mit einem Löffel die Suppe umzurühren? –, sodass es zu sehr speziellen Lösungen für eine Verhaltensaufgabe kommen kann, die weder robust noch auf andere Aufgaben oder andere Roboter übertragbar sind. Kleine Änderungen in der Umwelt oder der Kalibrierung des Roboters können dann komplettes Fehlverhalten zur Folge haben. Wenn heute Roboter in komplexen Domänen eingesetzt werden (sollen), dann ist es deshalb in der Regel so, dass diese nur semiautonom agieren: Der Roboter führt einzelne Aktionen autonom aus, die übergeordnete Steuerung sowie Zielfestlegung ist jedoch weiterhin Sache des

74 Ein anderes Lernparadigma, das sich stärker am menschlichen Lernverhalten anlehnt, ist das Lernen aus Vorbildern. Dabei werden komplexe Handlungen anhand weniger Beispiele durch Vormachen oder Imitation auf den Roboter übertragen. Voraussetzung dafür ist, dass zwischen Roboter und Vorbild eine hinreichende Ähnlichkeit besteht, damit Beispielhandlungen auf dem Roboter auch in geeigneter Weise repräsentiert werden können (Behnke 2014, S.13).



Menschen. Beispiele für eine enge Aufgabenverteilung zwischen Mensch und Roboter finden sich in Fahrerassistenzsystemen oder teleoperierten Robotern. Auch die kürzlich abgeschlossene »DARPA Robotics Challenge«,⁷⁵ in der Roboter komplexe Aufgaben in Katastrophenszenarien lösen sollten, erlaubte die Fernsteuerung der Roboter – jedoch unter der Bedingung unzuverlässiger Kommunikation, sodass die Teams zumindest gezwungen waren, einen gewissen Grad an Autonomie zu verwirklichen (»supervised autonomy«). Entscheidend für den Erfolg einer derartigen Mensch-Roboter-Kooperation ist die Gestaltung der Benutzerschnittstelle. Wie kann der Mensch die Situation des Roboters möglichst intuitiv erfassen und möglichst einfach mit dem Roboter zusammenarbeiten? Auch der Wechsel zwischen Autonomiegraden, wie beispielsweise die Übergabe der Kontrolle von einer an ihre Grenzen stoßenden autonomen Fahrzeugsteuerung an den Menschen, muss noch weiter untersucht werden.

Systemintegration

Autonome Roboter sind mehr als die Summe ihrer Teile (dazu und zum Folgenden Behnke 2014, S.14, sowie Schaal 2014, S.36). Als Perception-Action-Learning-Systeme operieren ihre einzelnen Komponenten nicht unabhängig voneinander, sondern in engster Interaktion. Somit entsteht erst durch die mechanische, elektrische und informationstechnische Integration der verschiedenen Systembestandteile in einem kohärenten Gesamtsystem in der Summe sinnvolles Roboterverhalten. Passende und qualitativ hochwertige Hardware – insbesondere standardisierte Schnittstellen (z. B. Controller Area Network [CAN] oder EtherCAT), aber auch leistungsfähige Sensoren – ist dafür notwendig, aber keineswegs ausreichend. Ein wesentlicher Teil der Funktionalität eines Roboters wird durch Softwaresysteme hergestellt, und genau diese Softwaresysteme stellen einen der Knackpunkte für die autonome Robotik dar.

Auch die Robotersoftware besteht aus zahlreichen Komponenten, die in der Regel auf mehreren Steuercomputern laufen, unterschiedliche Funktionalitäten des Systems bereitstellen (Wahrnehmungs-, Steuerungs- und Lernalgorithmen) und durch spezielle Programme (sogenannte Middleware) verbunden werden müssen (dazu und zum Folgenden Behnke 2014, S.14 f., sowie Schaal 2014, S.37). Bislang ist es so, dass nahezu jedes Labor hierfür eigene Lösungen entwickelt, die oft nur auf einen einzigen Roboter und einige spezialisierte Anwendungen zugeschnitten sind und darüber hinaus auf ganz unterschiedlichen Methoden und Paradigmen der KI- und ML-Forschung beruhen können (z. B. Reinforcementlernen vs. Imitationslernen). Deswegen sind die Ergebnisse oft

⁷⁵ www.theroboticschallenge.org/ (15.12.2015)



nicht, oder nicht einfach, auf andere Systeme übertragbar. Versuche, generalisierte Softwaresysteme zu entwickeln, gibt es zwar, sie sind jedoch noch nicht ausgereift. Das Robot Operating System (ROS) von Willow Garage (USA), das von der Open Source Robotics Foundation und zahlreichen Forschergruppen quelloffen weiterentwickelt wird, ist weltweit vermutlich am meisten verbreitet. Der dahinter liegende Grundgedanke ist, eine Softwareplattform bereitzustellen, die unabhängig von spezieller Roboterhardware ist, viele Prozessmodule koordiniert, zudem einfach erweiterbar ist und es mithin erlaubt, relativ komplexe, aber dennoch standardisierte Softwarelösungen zu entwickeln. Viele Softwarekomponenten zur Wahrnehmung, Planung, Kommunikation, Überwachung und Visualisierung existieren bereits für ROS. Diese Module können auf speziellen Plattformen veröffentlicht und mit vergleichsweise geringem Aufwand an neue Roboter und Problemstellungen angepasst werden.

Obwohl ROS in vielen Forschungslaboren Anklang gefunden hat, ist es mit einigen Nachteilen verbunden: Zum einen unterstützt das System keine Echtzeitregelung, die bei vielen Robotern notwendig ist, zum anderen ist das Softwaresystem sehr komplex und zudem abhängig von dem Betriebssystem Linux, auf dem es läuft (dazu und zum Folgenden Schaal 2014, S. 37). Weiterhin ist die Qualität der verfügbaren Module oft mangelhaft – meistens handelt es sich dabei um schlecht dokumentierte Software, die schwer zu verstehen und oft nicht sehr zuverlässig ist. Benutzt man solche Software mit einem kostspieligen Roboter, besteht die Gefahr, dass ein Softwarefehler das System beschädigt. Neben ROS existieren viele weitere generalisierte Softwareentwicklungen für die Robotik, die jedoch alle wesentlich spezialisierter sind. So hat das Italian Institute of Technology (IIT) z. B. eine Softwareumgebung für den iCub-Roboter⁷⁶ geschaffen, die viele Funktionen für die Benutzer zur Verfügung stellt. Auch der französische Roboterhersteller Aldebaran Robotics⁷⁷ stellt eine ähnliche Softwareumgebung für seinen humanoiden Roboter NAO bereit. Zusätzlich gibt es spezialisierte Open-Source-Pakete für Computer Vision (OpenCV), Planen (RAVE), Echtzeitregelung (SL), Optimierung (DoCoMo) oder 3-D-Vision (PointCloud Library). Hierbei handelt es sich meist um kleinere Projekte, die aber oft von recht guter Qualität sind.

Letztendlich stellt die Softwaretechnik nach wie vor eine der wesentlichen Hürden bei der Entwicklung autonomer Roboter dar (dazu und zum Folgenden Schaal 2014, S. 37 f. u. 45). Es ist meistens relativ problemlos möglich, eine Robotikfunktion in einer Machbarkeitsstudie so zu programmieren, dass das erwünschte Verhalten im Labor sporadisch demonstriert werden kann. Wesentlich schwieriger ist es jedoch, ein robustes autonomes System mit reproduzierbarem Verhalten herzustellen, dessen Software optimal auf die Hardware abge-

⁷⁶ www.icub.org/ (15.12.2015)

⁷⁷ www.aldebaran.com/en (15.12.2015)



stimmt ist. Wie groß diese Aufgabe ist, wurde auch in der bereits erwähnten DARPA Robotics Challenge deutlich. Der 2012 gestartete Wettbewerb hatte zum Ziel, die Entwicklung humanoider, semiautonomer Roboter zu fördern, die im Katastropheneinsatz nutzbar sein könnten. In einem Parcours mussten dazu, unter menschlicher Beihilfe, eine Reihe von Mobilitäts- und Manipulationsaufgaben bewältigt werden, wie z. B. das Überqueren eines Schuttfelds, das Klettern auf einer Leiter oder das Öffnen von Türen. Ein Video, in dem Ausschnitte aus dem Wettbewerb zu sehen sind, zeigt, dass viele der Roboter bereits an wesentlich geringeren Herausforderungen scheiterten und es nicht schafften, sich in ebenem Gelände fortzubewegen, ohne plötzlich umzufallen – und dies nach jahrelanger Entwicklungszeit.⁷⁸ Der Wettbewerb wurde 2015 abgeschlossen, gewonnen hat das Team des Korea Advanced Institute of Science and Technology mit dem humanoiden Roboter Hubo.

Das Bild der stolpernden Roboter der DARPA Robotics Challenge illustriert, dass in der Robotik ganz andere und konservativere Methoden des Programmierens erforderlich sind als bei reinen Software- oder KI-Systemen, da Softwarefehler wesentlich gravierendere Auswirkungen haben können (dazu und zum Folgenden Behnke 2014, S. 15, sowie Schaal 2014, S. 36 ff. u. 42). Trotz der Unterstützung durch Standards, Middleware und Softwarebibliotheken ist der dafür erforderliche Aufwand erheblich und erfordert interdisziplinäre Kompetenzen, über die nur wenige Softwarespezialisten verfügen. Aufgrund des großen Entwicklungsaufwandes sind gute Manipulationsroboter bislang sehr teuer – das Kostenspektrum für einen humanoiden Roboter beginnt normalerweise bei ca. 500.000 Euro und kann schnell bei mehreren Millionen Euro liegen. Daher ist für Forschungsaufgaben der Einsatz von standardisierten Plattformen – einheitlicher Hardware und Systemsoftware – der Entwicklung eines eigenen Robotersystems häufig vorzuziehen. Einer der bedeutendsten Forschungsroboter ist iCub (Abb. III.13), der 2004 im Rahmen des europäischen RobotCup-Projekts unter Führung des Italian Institute of Technology (IIT), entstanden ist.⁷⁹ Mit 53 Gelenken von beachtlicher Komplexität und von Größe und Erscheinung her einem 4-jährigen Kind nachempfunden, wurde iCub entwickelt, um grundlegende Fragen zur Entwicklung von Motorik, Wahrnehmung und Lernkompetenz zu erforschen (Metta et al. 2008). Mittlerweile wird der Roboter von ca. 20 Forschungslabors im Rahmen verschiedener Grundlagenprojekte genutzt.⁸⁰ Andere wichtige offene Forschungsplattformen sind der Roboter NAO der Firma Aldebaran,⁸¹ der Roboter PR2 von Willow

78 www.youtube.com/watch?v=g0TaYhjpOfo&feature=youtu.be. (15.12.2015)

79 www.icub.org/ (15.12.2015)

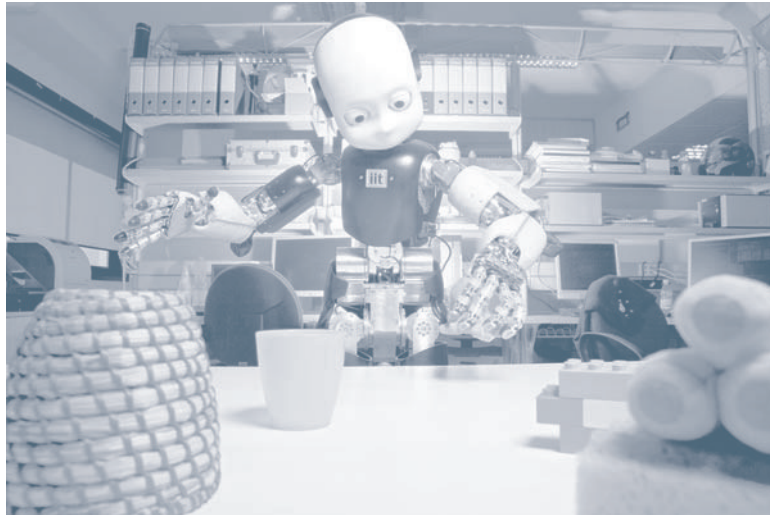
80 www.icub.org/projects.php (15.12.2015)

81 www.aldebaran.com/en/humanoid-robot/nao-robot (15.12.2015)

Garage⁸² sowie der Gewinner der DARPA Robotics Challenge, der Roboter Hubo des Korea Advanced Institute of Science and Technology (KAIST) (Park et al. 2007).

Abb. III.13

Der Roboter iCub



Quelle: Fondazione Istituto Italiano di Tecnologia

Anwendungsfelder und Beispiele

2.3

Übergeordneter Sinn und Zweck der autonomen Robotik ist die Entwicklung von Systemen, welche die Automatisierung menschlicher Tätigkeitsfelder vorantreiben. Ein besonderer Bedarf hierfür zeichnet sich insbesondere angesichts des demografischen Wandels ab, der die meisten hochentwickelten Staaten erfasst hat (dazu und zum Folgenden Schaal 2014, S. 19 ff.). Aufgrund sinkender Geburtsraten einerseits und besserer Ernährung und medizinischer Versorgung andererseits steigt der Anteil älterer und damit unterstützungs- und pflegebedürftiger Menschen an der Gesamtbevölkerung sukzessive an, während der Anteil der erwerbstätigen Bevölkerung entsprechend schrumpft. Das bedeutet, dass die Ressourcen für die Versorgung der alternden Bevölkerung von einer immer geringeren Zahl Menschen erwirtschaftet werden müssen. Angesichts dessen ist eine verbreitete Hoffnung für die Zukunft, dass autonome technische Helfer die zu erwartenden Produktivitätsverluste ausgleichen sowie auch den steigenden Betreuungs- und Versorgungsbedarf decken können. Das prognostizierte Marktpotenzial für autonome Roboter ist demnach enorm, und es er-

82 www.willowgarage.com/pages/pr2/overview (15.12.2015)



staunt nicht, dass sowohl private als auch öffentliche Akteure – zu nennen sind insbesondere die DARPA und die großen IT-Konzerne Google, Apple und Amazon – bereit sind, viel Geld in die autonome Robotik zu investieren.

Die Anwendungsmöglichkeiten für autonome Roboter sind im Einzelnen äußerst vielfältig und sprengen bei Weitem den rein demografisch begründeten Einsatzzweck:

- > Autonome Roboter sollen als Kampf-, Transport- oder Katastrophenroboter die menschlichen Aktionsmöglichkeiten erweitern, speziell in unzugänglichen oder gefährlichen Umgebungen;
- > sie sollen als Service- oder Assistenzroboter mühsame Alltagstätigkeiten im Haushalt erleichtern oder Dienstleistungen in Bereichen wie der Pflege oder dem Gastgewerbe erbringen;
- > eine ähnliche Bestimmung haben sogenannte soziale Roboter, eine spezielle Form der Serviceroboter, die mit einem menschlichen Gegenüber interagieren sollen, sei es als Spielgefährte, Therapeut oder Museumsführer;
- > und natürlich sollen sie auch als Industrieroboter dazu beitragen, die industrielle Produktivität zu steigern. Hier wird zunehmend an komplexere Aufgaben gedacht, die durchaus in enger Kooperation mit Menschen durchgeführt werden können – und nicht nur an monotone, hochspezialisierte Arbeiten in abgesperrten Bereichen, wie sie traditionelle Industrieroboter vollbringen;
- > ein weiteres, jedoch noch rein futuristisches Anwendungsszenario ist der medizinische Einsatz als Nanoroboter, die sich im Körperinnern über die Blutbahn fortbewegen, um körperliche Prozesse zu überwachen oder in diese einzugreifen.

Die technologischen Anforderungen sind im Einzelnen so unterschiedlich wie die Anwendungsfelder selber. Um ein genaueres Bild der heterogenen FuE-Dynamik zu bekommen, werden im Folgenden exemplarische Anwendungen aus den einzelnen Einsatzbereichen vorgestellt und ihr Entwicklungsstand sowie die sich dabei stellenden Herausforderungen und Lösungsansätze kursorisch diskutiert.

Autonome Roboter für industrielle Anwendungen

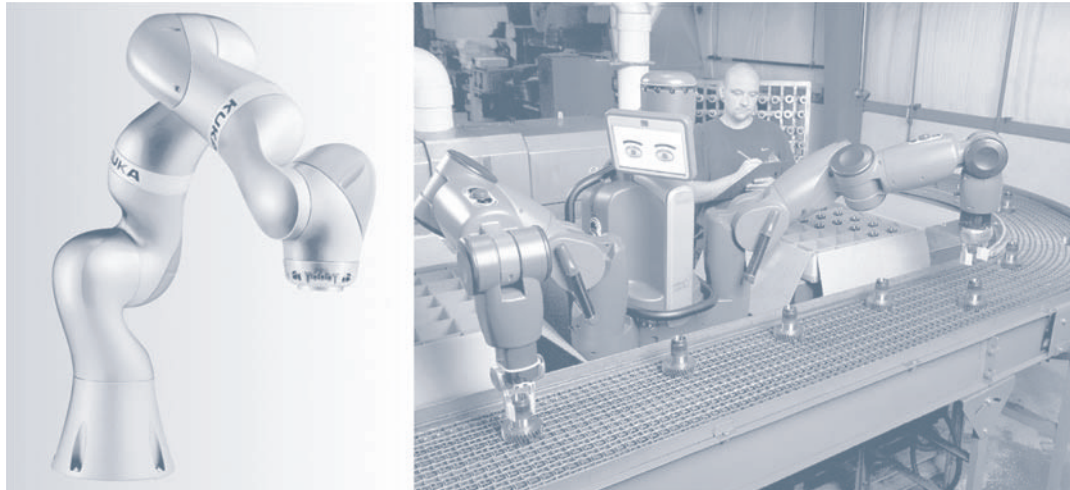
Wie bereits erwähnt, verfügen herkömmliche Industrieroboter nur über minimale Sensorik und kaum über autonome Fähigkeiten. Sie werden deshalb ausschließlich in abgesperrten Bereichen und für repetitive Tätigkeiten in einem Produktionssystem eingesetzt, das ganz auf sie abgestimmt sein muss. Die Robotik erobert aber auch zunehmend neue industrielle Bereiche, eine Entwicklung, die in Deutschland unter dem Schlagwort Industrie 4.0 diskutiert wird

(Hirsch-Kreinsen/Weyer 2014). So taucht am Horizont eine neue Generation von autonomeren Industrierobotern auf, die mit Menschen kollaborieren, komplexe Aufgaben selbst erfüllen und sich flexibel an neue Aufgaben anpassen können. Diese Systeme sind v. a. aus Sicherheitsgründen wesentlich leichter gebaut als klassische Industrieroboter, außerdem mit nachgiebiger Steuerung und einer hohen Anzahl diverser Sensoren ausgestattet (Schaal 2014, S. 27).

Ein typisches Beispiel für diesen neuen Typus von Industrieroboter ist der Leichtbauroboter der Firma Kuka, der ursprünglich in den Forschungslaboren des Deutschen Zentrums für Luft- und Raumfahrt (DLR) entwickelt wurde, einem der in Europa führenden Institute für Mechatronik und zukunftsorientierte Robotik (Abb. III.14 links) (Bischoff et al. 2010).

Abb. III.14

Leichtbauroboter LBR iiwa und Baxter



Quelle: KUKA Roboter GmbH (links); Rethink Robotics, Inc. (rechts)

Es handelt sich hierbei im Wesentlichen um einen flexiblen Roboterarm, der dank seiner ausgefeilten Sensorik als »äußerst feinfühlig« gilt und aufgrund seiner sieben Achsen über eine Beweglichkeit verfügt, die sogar die eines menschlichen Arms übertreffen soll.⁸³ Er eignet sich damit vor allem für feinmechanische Handhabungs- und Manipulationsaufgaben, die bislang ohne menschliches Fingerspitzengefühl nicht durchführbar waren, etwa Schweißen oder komplizierte, kleinteilige Montagevorgänge. Dank seiner modularen Bauweise und seines geringen Gewichts von etwa 20 kg lässt er sich flexibel für vielfältige Aufgaben einsetzen – nicht nur im industriellen, sondern etwa auch im medizinischen Bereich. Mit dem neuesten Modell des Leichtbauroboters, dem LBR iiwa,

83 www.kuka-lbr-iiwa.com/de/ (15.12.2015)



hat Kuka erstmals einen serienreifen Leichtbauroboter für den industriellen Dauereinsatz vorgestellt.

Stärker an mittelständische Industrieunternehmen richtet sich der kollaborative Roboter Baxter der Firma Rethink Robotics, der über zwei Manipulationsarme und humanoides Aussehen verfügt (Abb. III.14 rechts).⁸⁴ Der ca. 30.000 Euro teure Baxter ist dazu gedacht, einfachere Arbeitsabläufe flexibel zu automatisieren, ohne dass dafür spezielle Sicherheitsmaßnahmen erforderlich wären (dazu und zum Folgenden Schaal 2014, S. 31). Baxter lernt durch Vormachen der Aufgaben, womit er sich unkompliziert in neue Prozesse integrieren lassen soll. Von der Funktionalität her kommt er jedoch bei Weitem nicht an die deutlich teureren Roboterarme von Kuka heran (ein Kuka Leichtbauroboter kostet ca. 100.000 Euro), weshalb das Konzept bei Unternehmen anscheinend auch noch keinen großen Anklang gefunden hat.

Während leichte, flexibel einsetzbare Manipulationsroboter, wie der Kuka Leichtbauroboter und Baxter, noch Exoten in der Industriewelt sind, haben autonome Transportsysteme auf Rädern schon eine deutlich weitere Verbreitung gefunden. Ein Beispiel sind etwa die mobilen Lagerroboter, die von der Firma Kiva Systems entwickelt wurden und die der Automatisierung von Warenlagern dienen (dazu und zum Folgenden Behnke 2014, S. 20). Das System besteht aus zahlreichen mobilen Robotern, die sich teilautonom in Lagerhallen bewegen können und Transportaufträge erledigen.⁸⁵ Amazon, das laut eigenen Angaben etwa 15.000 solcher Roboter im Einsatz hat,⁸⁶ hat Kiva Systems 2012 für 775 Mio. US-Dollar erworben und in Amazon Robotics umbenannt. Ein ähnliches Beispiel sind die fahrerlosen Transportsysteme der Firma swisslog,⁸⁷ die für Hol- und Bringdienste im Krankenhaus entwickelt wurden. Um Hindernisse zuverlässig zu erkennen und Kollisionen zu vermeiden, sind sie mit einem 270°-Sicherheitslaserscanner ausgestattet, der gleichzeitig der Lokalisierung dient. Die Roboter stoppen vor Hindernissen, planen aber keine Alternativroute, wenn der vordefinierte Pfad dauerhaft blockiert ist.

Was in semistrukturierten Gebieten, wie z. B. in einem Warenlager, schon recht gut funktioniert, stößt in unstrukturierteren Umgebungen, etwa einem

⁸⁴ www.rethinkrobotics.com/baxter/ (15.12.2015)

⁸⁵ Jeder Roboter kann ein Regal mit Waren anheben und transportieren (dazu und zum Folgenden Behnke 2014, S. 20). Ein Managementsystem sorgt dafür, dass die zu einem Packauftrag gehörenden Regale in der richtigen Reihenfolge zu einer Packstation gebracht werden, an der ein Mensch das Greifen und Verpacken der Ware erledigt. Die Regale werden so im Lager verteilt, dass häufig benötigte Waren die kürzesten Wege haben.

⁸⁶ <http://phx.corporate-ir.net/phoenix.zhtml?c=176060&p=irol-newsArticle&ID=1993497> (15.12.2015)

⁸⁷ www.swisslog.com/de/Products/HCS/Automated-Material-Transport/TransCar-Automated-Guided-Vehicles (15.12.2015)



Krankenhaus, noch auf größere Schwierigkeiten. Insbesondere wird es verstärkter Forschungsanstrengungen hinsichtlich Wahrnehmung, Regelung und Anpassungsfähigkeit bedürfen, bis autonome Manipulationsroboter in industriellen Kontexten wirtschaftlich einsetzbar sind. Dies zeigte auch die Amazon Picking Challenge, die kürzlich mit dem Ziel veranstaltet wurde, das automatisierte Greifen von Gegenständen zu verbessern.⁸⁸ Aufgabe des Wettbewerbs war es, einen Roboter bzw. Roboterarm zu entwickeln, der verschiedene kleinere Alltagsgegenstände in einem Regal identifizieren und sicher daraus entnehmen kann. Auch der Roboter des siegreichen Teams RBO von der TU Berlin erledigte diese Aufgaben nur im Zeitlupentempo, wie auf einem Internetvideo zu sehen ist,⁸⁹ und war nicht imstande, in 20 Minuten alle zwölf Objekte fehlerfrei abzulegen.

Roboter für unzugängliche oder gefährliche Umgebungen

Der Einsatz von autonomen Robotern ist besonders dann angezeigt, wenn Aufgaben in Situationen erledigt werden müssen, die für Menschen zu gefährlich oder nicht zugänglich sind (dazu und zum Folgenden Behnke 2014, S. 24, sowie Schaal 2014, S. 15). Beispiele dafür sind Katastrophen, der Weltraum oder Kriege. Bereits in den 1950er Jahren gab es die ersten aus der Distanz gesteuerten Teleroboter, etwa um radioaktive Gegenstände aus der Entfernung zu manipulieren. Diese Systeme waren jedoch weder autonom, noch entsprachen sie dem modernen Bild eines Roboters: Es handelte sich im Wesentlichen um rein mechanische Robotikarme, die über ein externes Handhabungssystem gesteuert wurden. Heute werden im Idealfall, wie es zum Beispiel bei modernen Drohnen bereits der Fall ist, autonome oder semiautonome Roboter verwendet, die wesentlich komplexere Aufgaben (Rettungseinsätze, Aufräum- und Reparaturarbeiten, Überwachung und Aufklärung) in unstrukturierten Umgebungen selbst durchführen können und dabei nur punktuell externe Hilfe über eine Fernsteuerung benötigen.⁹⁰

Die Anforderungen an Fortbewegung, Manipulation und Aufklärung sind im Katastropheneinsatz besonders hoch (dazu und zum Folgenden Behnke

88 <http://amazonpickingchallenge.org/> (15.12.2015)

89 www.youtube.com/watch?v=UrpMfdj-Mpc&feature=youtu.be (15.12.2015)

90 Beim Einsatz in für Menschen unzugänglichen Umgebungen spielen die durch Teleoperation entstehenden Personalkosten nur eine untergeordnete Rolle (dazu und zum Folgenden Behnke 2014, S. 25). Bei guter Kommunikationsverbindung zwischen Roboter und Operator erscheint daher oft eine direkte Kontrolle des Roboters wünschenswert. Dennoch werden teilautonome Verhaltensweisen entwickelt, um die Aufgaben schneller und präziser zu erfüllen und auch bei fehlender Kontrolle die Einsatzfähigkeit sicherzustellen.



2014, S.24; Schaal 2014, sowie S.2 u. 16).⁹¹ Dass die zur Verfügung stehenden Technologien diese Ansprüche bei Weitem noch nicht erfüllen können, zeigt die bereits erwähnte DARPA Robotics Challenge wie auch die Reaktorkatastrophe von Fukushima. Obwohl Japan in der Robotik zu den führenden Nationen zählt, stand kein autonomes System bereit, das sinnvoll Hilfe hätte leisten können. Bei den zur Verfügung stehenden Technologien handelt es sich um mobile Roboter (auf Rädern), die – meist ferngesteuert – zur Aufklärung, zur Bergung gefährlicher Gegenstände bzw. zur Entschärfung von Sprengsätzen eingesetzt werden. Abbildung III.15 (links) zeigt den Roboter Packbot⁹² der Firma iRobot im Einsatz. Dieser Roboter kann mit einem 3-D-Laserscanner die Einsatzumgebung erfassen, Karten erzeugen und autonom navigieren. Wesentliche Impulse zur Entwicklung autonomer Robotersysteme kommen aus der Weltraumforschung, wo eine Teleoperation aufgrund der langen Signallaufzeit sowie der begrenzten Kommunikationsmöglichkeiten naturgemäß an Grenzen stößt. Im Rahmen des NASA Robonaut-Projekts⁹³ wird seit vielen Jahren an Robotern dieser Art geforscht. Der Curiosity-Rover⁹⁴ der NASA befindet sich seit 2012 auf dem Mars und führt dort mit zehn Instrumenten Untersuchungen durch. Der Roboter kann autonom navigieren, wird aber vom NASA Jet Propulsion Laboratory aus konfiguriert und überwacht.

Ein besonders umstrittenes Gebiet ist der Einsatz militärischer Roboter (dazu und zum Folgenden Schaal 2014, S.17 u. 40). Einerseits möchte man Menschen aus militärischen Gefahrenzonen möglichst fern halten, andererseits wirft die zunehmende Automatisierung von Kampfhandlungen viele ethische Problematiken auf, die aktuell kontrovers diskutiert werden (International Committee of the Red Cross 2014) – in einem offenen Brief wurde, wie bereits erwähnt (Kap. II.3), jüngst von über 20.000 Personen, darunter viele KI-Forscher und Berühmtheiten wie Stephen Hawking und Elon Musk, ein Verbot offensiver autonomer Waffen gefordert.⁹⁵ Im Vordergrund der Debatte steht die Frage, inwiefern es ethisch vertretbar und völkerrechtlich zulässig ist, Maschinen im Gefecht autonom über Leben und Tod von Menschen entscheiden zu lassen (Royakkers/van Est 2015, S.249 ff.). Bereits heute befinden sich diverse, roboterähnliche Maschinen im militärischen Einsatz, ihre Autonomie ist allerdings noch eher begrenzt. Sicherlich am weitesten verbreitet sind Drohnen, die zwar bereits in großer Zahl militärische Handlungen ausführen (TAB 2011), dabei

91 www.wissenschaft.de/technik-kommunikation/technik/-/journal_content/56/12054/939663/Kein-Markt-f%C3%BCr-Katastrophen-Roboter/ (15.12.2015)

92 www.irobot.com/For-Defense-and-Security/Robots/510-PackBot.aspx#PublicSafety (15.12.2015)

93 <http://robonaut.jsc.nasa.gov/default.asp> (15.12.2015)

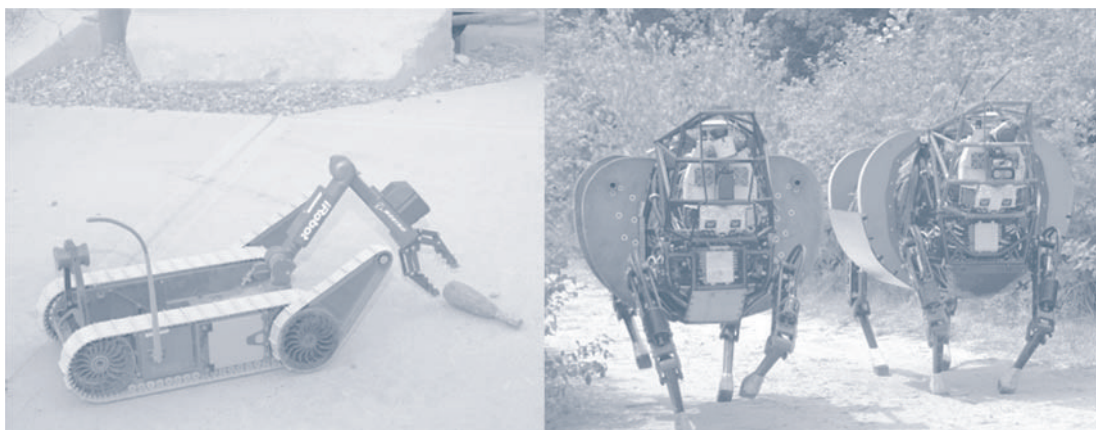
94 www.nasa.gov/mission_pages/msl/ (15.12.2015)

95 http://futureoflife.org/AI/open_letter_autonomous_weapons (15.12.2015)

jedoch ferngesteuert werden, sodass sich ihre Selbstständigkeit im Wesentlichen auf die stabile Flugkontrolle beschränkt. Ebenfalls einen stark militärischen Anwendungsbezug, wenngleich keinen Waffencharakter, haben der bereits erwähnte Packbot, der seit 2002 in tausendfacher Zahl von der US-Armee u. a. als Spähroboter eingesetzt wird, sowie die Transportroboter »BigDog« und dessen Nachfolger »LS3« (Legged Squad Support Systems)⁹⁶ (Abb. III.15 rechts). Hierbei handelt es sich um einen maultierähnlichen Lastenroboter auf vier Beinen, der über einen Benzinmotor angetrieben wird und bis zu 180 kg transportieren kann. Das von der Firma Boston Dynamics⁹⁷ entwickelte Projekt wurde 2005 von der DARPA beauftragt und mit ca. 100 Mio. US-Dollar unterstützt, erste Prototypen werden von der US-Armee getestet. LS3 kann sich auch in schwierigem Gelände sicher und semiautonom fortbewegen, dabei wird er über Fernsteuerung dirigiert oder er wird darauf programmiert, mithilfe seiner visuellen Sensoren einem Menschen zu folgen.

Abb. III.15

Packbot und LS3



Quelle: U.S. Army (links); DARPA (rechts)

Es wird erwartet, dass sich derartige Roboter »in den kommenden Jahren ... von reinen ferngesteuerten Maschinen – etwa zur Bombenentschärfung oder als bewegliche Videokamera in Gebäuden – zu vernetzten Systemen fortentwickeln, die eigenständig agieren und Aufträge erfüllen« (Dickow 2015, S. 17). Die ethischen, rechtlichen und sicherheitspolitischen Fragen, die diese Entwicklung im Hinblick auf mögliche Kampfeinsätze aufwirft, sind drängend. Gleichwohl ist festzuhalten, dass noch mindestens 10 bis 20 Jahre vergehen dürften, bis komplett autonome Roboter in unstrukturiertem Gelände wirklich einsatzfähig

⁹⁶ www.bostondynamics.com/robot_ls3.html (15.12.2015)

⁹⁷ www.bostondynamics.com (15.12.2015)



sind (Schaal 2014, S. 72). Die Autonomie, welche die Erledigung etwa spezieller militärischer Aufgaben am Boden erfordert, liegt noch weit jenseits dessen, wozu die Technologien derzeit imstande sind. Bei Drohnen dürfte sich die Entwicklung in Richtung Autonomie hingegen deutlich schneller vollziehen.

Serviceroboter

Der Servicebereich ist neben der Industrie sicherlich eines der zukunftsreichsten Einsatzgebiete für autonome Roboter. Laut International Federation of Robotics (IFR) sind Serviceroboter gerade dabei, die Welt zu erobern: Seit Jahren steigen die Verkaufsraten im zweistelligen Bereich (IFR 2015b).⁹⁸ Für einfache Aufgaben in Haushalt und Garten, wie Staubsaugen und Rasenmähen, sind kostengünstige Serviceroboter schon heute verfügbar und haben eine weite Verbreitung gefunden (dazu und zum Folgenden Behnke 2014, S. 19). Beispielsweise bietet die Firma iRobot Staubsaugerroboter, Bodenreinigungsroboter, Bodenwischroboter sowie Roboter zur Reinigung von Swimmingpools und Dachrinnen an. Diese Geräte – die Bezeichnung Roboter ist hier sicherlich hoch gegriffen – sind mit einfachen Sensoren wie Abstandssensoren und Kontaktsensoren ausgestattet und haben eine geringe Rechenleistung. Typisch ist das quasi zufällige Fahren bis zu einem Hindernis, an dem die Richtung geändert wird. Wenn die Umgebung robotergerecht gestaltet wird, z. B. durch Vermeiden auf dem Boden liegender Kabel, können sie ihre Aufgaben selbstständig erledigen. Sie finden auch ihre Ladestation wieder, die im Infrarotbereich Signale aussendet. Auch einfache Rasenmäherroboter verfolgen eine ähnliche Strategie. Dazu wird die zu mähende Fläche in der Regel durch ein Kabel eingegrenzt. Damit die Roboter systematischer vorgehen können, werden bessere Sensoren und mehr Rechenleistung benötigt.⁹⁹

Entscheidend für den Markterfolg solcher Helfer ist, dass der Mehrwert, der durch Arbeitserleichterung oder erhöhten Komfort entsteht, die Anschaffungs-

98 So stieg der weltweite Verkauf von Haushaltsrobotern um 28 % auf ca. 3 Mio. Einheiten. Bei den Servicerobotern für professionelle Anwendungen wurden 2014 rund 24.200 Einheiten verkauft (v. a. in den Bereichen Verteidigung und Landwirtschaft, dort insbesondere Melkroboter), was einer Steigerung gegenüber 2013 von 11,5 % entspricht. Als Zukunftsmarkt gelten private Assistenzroboter für die Senioren- und Behindertenbetreuung, die einen massiven Zuwachs zu verzeichnen haben (2014: 4.416 verkaufte Einheiten gegenüber 699 Einheiten in 2013) (IFR 2015b). Darauf hinzuweisen ist, dass die IFR den Begriff Serviceroboter sehr weit definiert und ex negativo alle Roboter darunter fasst, die nicht zu den industriellen Produktionsrobotern zählen.

99 Kürzlich hat die Firma Dyson einen Staubsaugerroboter auf den Markt gebracht, der mit einer Rundumkamera ausgestattet ist, eine Karte der Einsatzumgebung erstellt, sich lokalisiert und systematisch die zu reinigende Fläche abdeckt. Ähnliche Ziele verfolgt Bosch mit dem LogiCut-Navigationssystem für Rasenmäherroboter.



kosten rechtfertigt (Behnke 2014, S.19). Dies scheint im Großen und Ganzen bereits der Fall zu sein, wie die rasant wachsenden Verkaufszahlen andeuten. 2014 wurden laut IFR weltweit 4,7 Mio. Haushaltsroboter verkauft, 28 % mehr als 2013 (IFR 2015b). Geschätzt wird, dass 2015 bis 2018 noch einmal 25 Mio. privat genutzte Geräte hinzukommen. Trotz dieser Erfolgswahlen ist die Autonomie der bislang verfügbaren Serviceroboter bescheiden. In der Regel bewegen sie sich auf Rädern fort und können somit nur in begrenztem Umfang Hindernisse überwinden, außerdem verfügen sie nur über geringe Manipulationsmöglichkeiten – sie eignen sich folglich nur für hochspezialisierte, repetitive Aufgaben in klar geordneten (sogenannten robotergerechten) Arbeitsumgebungen. Indes wird in Forschung und Entwicklung bereits seit Längerem intensiv an einer neuen Generation von Servicerobotern gearbeitet, die fähig sein sollen, unterschiedliche Aufgaben auch in unstrukturierten Alltagsumgebungen selbstständig zu bewältigen. Dazu bedarf es hochentwickelter Interaktionsmöglichkeiten sowie der Fähigkeit, sich mittels zahlreicher Sensoren auch in veränderten Umgebungen zurechtzufinden und aus Misserfolgen zu lernen (dazu und zum Folgenden Behnke 2014, S.23 f.). Das Anwendungs- und Marktpotenzial für autonome Helfer, die über diese Kompetenzen verfügen, ist sehr groß. Sie könnten es hilfsbedürftigen Menschen ermöglichen, länger selbstbestimmt im gewohnten Umfeld zu leben, und auch in professionellen Pflegeeinrichtungen zahlreiche Aufgaben übernehmen. Große Potenziale bietet auch die Serviceindustrie, in der es viele scheinbar einfache Tätigkeiten gibt, die Roboter erledigen könnten, wie das Säubern von Tischen im Restaurant oder die Herrichtung von Gästezimmern im Hotel. Der größte Markt ist sicher der breite Einsatz im Haushalt, um jedermann von Alltagstätigkeiten zu entlasten.

Noch sind die Früchte der langjährigen Entwicklungsanstrengungen aber nicht ausgereift. Die vorliegenden Prototypen – Beispiele sind »PR2«¹⁰⁰ (Willow Garage), »Twendy-One«¹⁰¹ (Waseda Universität), »Care-O-Bot«¹⁰² (Fraunhofer IPA), »Rollin' Justin61«¹⁰³ (DLR), »Armar«¹⁰⁴ (Karlsruher Institut für Technologie) oder »Cosero«¹⁰⁵ (Universität Bonn) – haben einen deutlich stärkeren Forschungs- als Anwendungsbezug und sind noch sehr teuer (die Kosten gehen in die Hunderttausende Euro) (dazu und zum Folgenden Behnke 2014, S.23). Auffällig ist die verbreitete humanoide Gestalt, die es u. a. ermöglichen soll, in einem Umfeld zu agieren, das an den Menschen angepasst ist. Das bereits bestehende Potenzial persönlicher Serviceroboter wird bei Demonstrationen deut-

100 www.willowgarage.com/pages/pr2/overview (15.12.2015)

101 <http://twendyone.com/> (15.12.2015)

102 www.care-o-bot.de/de/care-o-bot-4.html (15.12.2015)

103 www.dlr.de/rm/desktopdefault.aspx/tabid-5471 (15.12.2015)

104 <http://his.anthropomatik.kit.edu/241.php> (15.12.2015)

105 www.ais.uni-bonn.de/nimbro/@Home/ (15.12.2015)

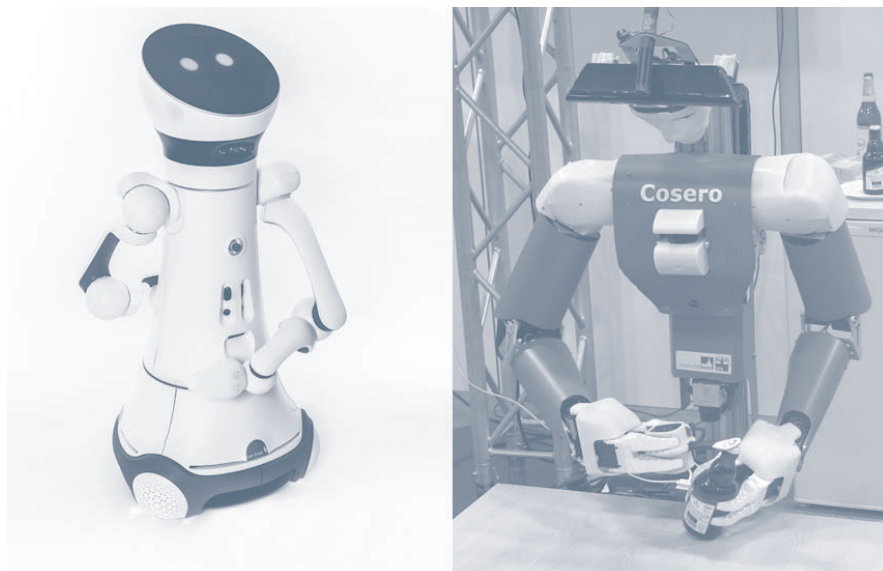


lich, bei denen diese komplexe Aufgaben lösen, wie das Einräumen von Geschirr in eine Spülmaschine oder das Backen von Pfannkuchen (Beetz et al. 2011). Allerdings erfordern diese Demonstrationen in der Regel immer noch viel Entwicklungsaufwand und funktionieren nur unter hinreichend kontrollierten Bedingungen. An einen produktiven Einsatz unter realen Bedingungen ist somit noch lange nicht zu denken.

Eines der am weitesten entwickelten Systeme, das der Marktreife schon recht nahe kommt, ist der vom Fraunhofer-Institut für Produktionstechnik und Automatisierung entwickelte Care-O-Bot, der seit 2015 in der vierten Generation vorliegt (Abb. III.16 links). Er wird als »elektronischer Butler« angepriesen, der »so zuvorkommend, freundlich und sympathisch wie ein Gentleman« sei (Fraunhofer-Institut für Produktionstechnik und Automatisierung 2015). Der Roboter, der sich dank innovativer Kugelgelenke in Hüfte und Hals sogar bücken kann und einfache Gesten wie Nicken oder Kopfschütteln beherrscht, ist modular aufgebaut und lässt sich somit für vielseitige Einsatzzwecke konfigurieren. Der ursprünglich intendierte Anwendungszweck ist, das lässt sich bereits dem Namen entnehmen, der pflegerische Einsatz zuhause oder in einer professionellen Einrichtung. Ein typisches Einsatzszenario, das auch schon real getestet wurde, ist das Holen oder Bringen von Gegenständen, damit sich die Pfleger auf ihre eigentliche Aufgabe konzentrieren können. Die Bedienung erfolgt über einen am Kopf angebrachten Touchscreen, zusätzlich verfügt das System über Mikrofone zur Spracherkennung und Kameras zur Personen- und Gestenerkennung.

Abb. III.16

Care-O-Bot 4 und Cosero



Quelle: Fraunhofer IPA (Foto: Rainer Bez) (links); Universität Bonn (rechts)

Um die Forschung im Bereich komplexer Serviceroboter für häusliche Anwendungen zu fördern und unterschiedliche Ansätze vergleichbar zu machen, veranstaltet die RoboCup-Federation jährliche internationale Wettbewerbe in der @Home-Liga (dazu und zum Folgenden Behnke 2014, S.23 f.; Stückler et al. 2012). Hier müssen die Roboter autonom eine Reihe komplexer Aufgaben lösen, wie beispielsweise das Holen von Getränken oder das Wegräumen herumliegender Gegenstände. Abbildung III.16 (rechts) zeigt den Serviceroboter Cosero der Universität Bonn im Einsatz, wie er eine Flasche zu öffnen versucht. Da die autonome Verhaltenssteuerung dieser Serviceroboter noch schnell an Grenzen stößt, wurde für Cosero ein Teleoperationsinterface auf einem Tablet-PC implementiert, mit dem ein Nutzer den Roboter aus der Ferne überwachen und konfigurieren kann (Schwarz et al. 2014). Die Steuerung erfolgt auf drei Ebenen: Auf der obersten Ebene spezifiziert der Nutzer Aufgaben, wie »hole das Bier aus dem Kühlschrank«. Auf der mittleren Ebene werden einzelne Aktionen konfiguriert: Der Roboter greift beispielsweise ein auf dem Bildschirm selektiertes Objekt oder fährt zu einer bestimmten Stelle. Auf der untersten Ebene kann der Nutzer den Roboter direkt steuern.

Ziel eines solchen Teleoperationsinterfaces ist es, den jeweils höchst möglichen Autonomiegrad zu nutzen, d. h., der Roboter soll möglichst viele Aufgaben selbst erledigen, um den Nutzer zu entlasten und auch um die Ausführungszeit niedrig zu halten (dazu und zum Folgenden Behnke 2014, S.24, sowie Schaal 2014, S.71). Auf diese Weise wird die Intelligenz des Nutzers mit der Fähigkeit zur mobilen Manipulation des Roboters kombiniert, was insbesondere für Nutzer mit Mobilitätseinschränkungen sehr hilfreich sein kann. Da im klinischen Bereich oft bereits eine vergleichsweise kleine Hilfe einen großen Unterschied machen kann und auch höhere Preise akzeptiert werden, könnten einfache Assistenzroboter dieser Art bereits in den nächsten 5 bis 10 Jahren von der Forschung in die breitere Anwendung übergehen.

Soziale Roboter

Die Robotik durchläuft momentan eine Entwicklung, die sich zunehmend von der rein industriellen Anwendung wegbewegt und stärker die Bereiche des normalen täglichen Lebens in den Blick nimmt (Schaal 2014, S.25). Dabei ist es fast zwangsläufig, dass Roboter und Menschen sich immer näherkommen, eine Entwicklung, die bereits bei den gängigen Servicerobotern zu beobachten ist. Für Service- und Industrieroboter wird zukünftig die intuitive Interaktion mit Menschen wesentlicher Teil der Funktionalität sein. Neben diesen etablierten Anwendungsbereichen eröffnen sich aber auch andere Einsatzfelder, bei denen die soziale Interaktion ganz zentral im Vordergrund steht: etwa therapeutische Aufgaben, die Betreuung von Kindern bzw. älteren Menschen, der Einsatz als



»Spielgefährte«, »Museumsführer« oder »Aufsichtsperson«. Roboter, die für solche Aufgaben geeignet sind, werden als soziale Interaktionsroboter bezeichnet – allerdings ist die Abgrenzung zum Serviceroboter dabei unscharf (Feil-Seifer/Matarić 2005). Bei sozialen Robotern handelt es sich im Grunde genommen um eine spezielle Ausprägung von Servicerobotern, bei denen die soziale Interaktionssituation den eigentlichen Anwendungszweck bestimmt und die deshalb über besondere Kommunikationsmittel sowie Ausdrucksmöglichkeiten verfügen müssen.

Kinder und ältere Menschen scheinen stark auf Interaktionsroboter zu reagieren, da sie eine ausgeprägte Veranlagung zur Imitation haben (dazu und zum Folgenden Behnke 2014, S.27 f.). In Pilotstudien konnte gezeigt werden, dass insbesondere autistische Kinder auf humanoide Robotern ansprechen und davon in ihrem Kommunikationsverhalten profitieren, wodurch sich neue Therapiemöglichkeiten ergeben (Wainer et al. 2010). Eine englische Forschergruppe hat spezifisch für den Umgang mit autistischen Kindern den humanoiden Roboter Kaspar entwickelt, der, vom Aussehen einer Spielzeugpuppe ähnlich, über sehr einfache gestische und mimische Ausdrucksmöglichkeiten verfügt.¹⁰⁶ Auf ähnliche Weise werden Roboter auch bereits bei älteren Menschen, etwa bei Demenzpatienten, zu therapeutischen Zwecken eingesetzt. Für einige Furore sorgte in diesem Zusammenhang die vom japanischen National Institute of Advanced Industrial Science and Technology (AIST) entwickelte Roboterrobbe »Paro« (Abb. III.17 links), die inzwischen auch in Deutschland zugelassen ist und in etwa 100 Altenheimen zum Einsatz kommt.¹⁰⁷ Paro ist als eine Art künstliches Haustier für demente Personen gedacht. Der etwa 50 cm große Roboter reagiert auf Berührung und Geräusche, er kann sich Personen zuwenden und ist fähig, Personen anhand ihrer Stimme zu unterscheiden. Das niedliche Äußere, die robbenähnlichen Geräusche und Bewegungen des Roboters wecken offenbar Fürsorgeinstinkte und sprechen die Demenzpatienten auf einer emotionalen Ebene an (Wada/Shibata 2007).

Ein anderes Beispiel für einen Interaktionsroboter ist »Robotinho«, der für den Einsatz als Museumsführer entwickelt wurde (dazu und zum Folgenden Behnke 2014, S.27; Nieuwenhuisen/Behnke 2013). Dieser Roboter kommuniziert mit den Besuchern über mehrere Kanäle gleichzeitig, verfügt über ausdrucksstarke Mimik und animiert den Mund beim Sprechen. Die Augen und Augenlider sind beweglich, und der Roboter erzeugt Blickkontakt mit seinen Kommunikationspartnern und wendet sich diesen mit Kopf und Rumpf zu. Der Robotinho ist fähig, verschiedene Gesten zu generieren – beispielsweise lenkt er die Aufmerksamkeit der Besucher durch Zeigegesten und gleichzeitiges Hinse-

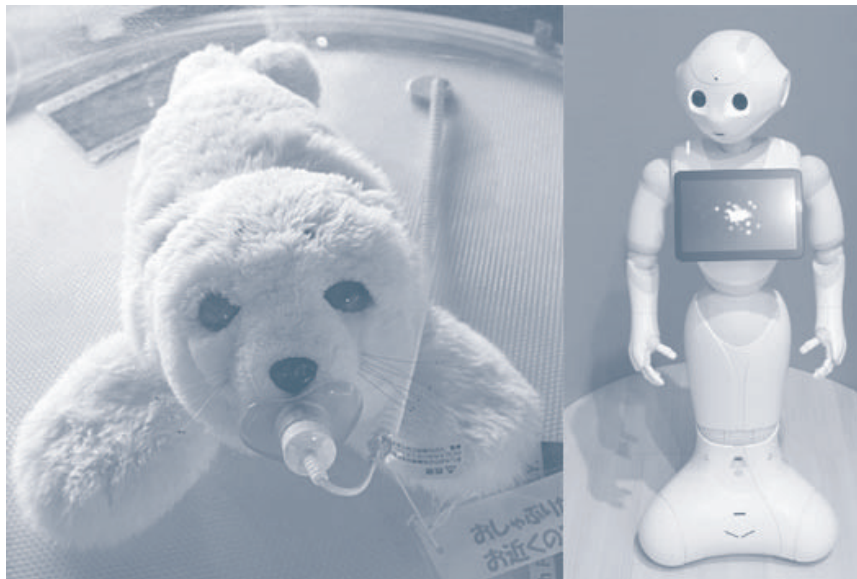
¹⁰⁶ www.herts.ac.uk/kaspar (15.12.2015)

¹⁰⁷ www.helmholtz.de/artikel/die-roboter-kommen-hilfe-in-der-pflege-und-im-alltag-2900/ (15.12.2015)

hen auf bestimmte Ausstellungsstücke, die er gerade erläutert. Auch bei der Navigation macht der Roboter durch seine Körpersprache klar, in welche Richtung er sich gerade bewegen möchte, bzw. behält die ihm folgenden Besucher im Blick. Dadurch, dass der Roboter die Kommunikationskanäle der zwischenmenschlichen Kommunikation nutzt, wirkt er besonders intuitiv. Robotinho wurde 2010 testweise im Deutschen Museum Bonn eingesetzt.

Abb. III.17

Paro und Pepper



Quelle: motoyen/Flickr/CC-BY-2.0 (links); Xavier Caré/Wikimedia Commons/CC-BY-SA-4.0 (rechts)

Es ist damit zu rechnen, dass im Zuge des demografischen Wandels zukünftig vor allem Zuwendungsroboter an Bedeutung gewinnen könnten, wenngleich deren pflegerischer Einsatz durchaus auf ethische Bedenken stößt (Christaller et al. 2001, S. 212). Bereits jetzt scheinen soziale Roboter ein recht großes Marktpotenzial zu haben (dazu und zum Folgenden Behnke 2014, S. 28, sowie Schaal 2014, S. 26 f.). Darauf deuten etwa die Vorverkaufszahlen für den einfachen Kommunikationsroboter »Jibo«¹⁰⁸ hin, der mehr als 4.800-mal vorbestellt wurde. Dabei kann der stationäre Jibo nicht viel mehr als ein modernes Smartphone – das jedoch zu einem recht günstigen Preis von ca. 750 US-Dollar. Er verfügt weder über Arme noch Beine, dafür jedoch über eine drehbare, kopf-ähnliche Kugel, in der ein Bildschirm und allerlei Technik (Raummikrofon, zwei HD-Kameras, Software zur Sprach- und Gesichtserkennung etc.) untergebracht sind. Damit soll er zu einfachen sozialen Interaktionen fähig sein, etwa

108 www.jibo.com/ (15.12.2015)



Personen am Gesicht zu erkennen und sie beim Namen zu nennen. Auch die französische Firma Aldebaran hat kürzlich mit Erfolg einen sozialen Gefährten – den humanoiden Familienroboter »Pepper«¹⁰⁹ (Abb. III.17 rechts) – auf den Markt gebracht. Auch hier waren die ersten 1.000 Exemplare sofort ausverkauft (zu einem Preis von je etwa 1.400 Euro). Pepper, der sich autonom fortbewegen kann und damit etwas komplexer aufgebaut ist als Jibo, verfügt ebenfalls über keine praktischen Fertigkeiten, sondern ist als lernfähiger Robotergefährte konzipiert, der auch auf emotionaler Ebene mit Menschen interagieren können soll. Neben dem privaten Bereich ist er deshalb primär für den Einsatz in Verkaufsräumen vorgesehen.

Bei diesen kommerziellen Beispielen scheint es sich aber vor allem noch um technische Spielereien zu handeln, von einem wirklichen Durchbruch in der sozialen Robotik kann noch keine Rede sein. Trotz teils sehr langer Entwicklungszeiträume – im Falle von Paro dauerte es von 1993 bis 2004, bis das Produkt in Serienreife vorlag – verfügen die beschriebenen Interaktionsroboter nur über minimale autonome Funktionen und nur über ein eingeschränktes Repertoire an Interaktionsmöglichkeiten. Während die Produktion von Sprache, Mimik, Gestik etc. teilweise schon recht gut funktioniert, ist die Erkennung der menschlichen Äußerungen, die für eine erfolgreiche soziale Interaktion wesentlich ist, häufig noch mit gravierenden Problemen behaftet und verbesserungsbedürftig. Einen autonomen sozialen Roboter mit komplexem Sozialverhalten zu entwickeln, stellt nicht zuletzt auch deshalb eine besondere, bislang noch nichtgelöste Herausforderung dar, weil es die Grenzen der herkömmlichen Robotikforschung sprengt (dazu und zum Folgenden Schaal 2014, S.26 f.). Zu den ursprünglichen Robotikthemen wie Lernen, Adaption, Planen, Regeltechnik etc. kommen neue Themengebiete hinzu wie das automatische Anpassen an verschiedene menschliche Verhaltensmuster, die Aufmerksamkeitskontrolle oder das Erkennen von Emotionen und Absichten (Schaal 2007). Nicht zu vergessen ist ein weiteres wichtiges Forschungsthema, nämlich das Design des Roboteransiehens, das entscheidenden Einfluss auf die Nutzerakzeptanz hat (Walters et al. 2008).

Nanoroboter

Ein Anwendungsfeld mit stark visionären Anklängen ist der Einsatz autonomer Nanoroboter im Inneren des menschlichen Körpers, die dort körperliche Funktionen selbstständig überwachen und diagnostische, therapeutische oder präventive Maßnahmen durchführen sollen (dazu und zum Folgenden Gransche/Warnke 2014, S.65 ff.). So futuristisch dies klingt: Die ursprüngliche Idee für

109 www.aldebaran.com/en/a-robots/who-is-pepper (15.12.2015)



»schluckbare Chirurgen« im Nanoformat stammt bereits aus den 1950er Jahren und wurde von dem Physiker Richard Feynman seinem Mitarbeiter Albert Hibbs zugeschrieben.¹¹⁰ Jahrzehnte danach wird zwar intensiv an mikroskopisch kleinen »Robotern« geforscht, die im menschlichen Körper eingesetzt werden sollen – und zwar durchaus mit einigem Erfolg (Tottori et al. 2014). Hierbei handelt es sich jedoch in der Regel um biologische Systeme auf bakterieller oder DNA-Basis, die aus technischer Sicht wenig bis gar nichts mit dem hier verwendeten Roboterbegriff zu tun haben, obwohl sie Nanoroboter genannt werden (Mavroidis/Ferreira 2013).

Gleichwohl wird – an der Schnittstelle Robotik, Materialforschung, Biochemie und Biophysik – weiterhin auch an einer stärker technisch orientierten Umsetzung der ursprünglichen Vision gearbeitet (dazu und zum Folgenden Schaal 2014, S.23 f. u. 72). Die Hindernisse für die Entwicklung von Nanorobotern im eigentlichen Sinne sind jedoch sehr groß, da im Mikro- und Nanometerbereich ganz andere physikalische Gesetzmäßigkeiten herrschen, etwa in Bezug auf den Flüssigkeitswiderstand, als wir sie aus alltäglichen Größenordnungen kennen (Spatz/Schaal 2014). Herkömmliche Konzepte und Ansätze aus der Robotik helfen hier nicht weiter. Folglich ist die Wissenschaft derzeit noch mit sehr grundlegenden Fragen befasst und versucht zu verstehen, wie man mit sensorischen und motorischen Funktionen ausgestattete Nanoroboter überhaupt generieren kann und wie es gelingen könnte, Informationsverarbeitung und Energieversorgung in diesen kleinen Systemen zu realisieren. Derartige Probleme werden die Grundlagenforschung voraussichtlich noch für die nächsten 10 bis 30 Jahre beschäftigen. Dennoch ist es nicht ausgeschlossen, dass schon früher in einigen klinischen Bereichen biobasierte Nanoroboter eingesetzt werden, etwa um Arzneimittel gezielt in den Körper einzubringen. Die Risiken solcher Praktiken sind jedoch noch völlig unklar, genauso wie man bisher nicht weiß, wie man Nanoroboter wieder aus dem Körper extrahieren könnte.

Forschungslandschaft in Deutschland

2.4

Vor allem aufgrund seiner starken Automobilindustrie zählt Deutschland, gemessen an den Verkäufen im Jahr 2014, zum weltweit fünftgrößten Markt für Industrieroboter (nach Japan, China, den USA und Korea) (IFR 2015a). Hinsichtlich der Roboterdichte steht Deutschland (mit 282 installierten Industrierobotern pro 10.000 Beschäftigte in der Produktionsindustrie) sogar auf dem dritten Platz nach Korea und Japan. Diese wettbewerbsstarke Position im Bereich der Industrierobotik entspricht den wissenschaftlich-technischen und

¹¹⁰ www.zyvex.com/nanotech/feynman.html (15.12.2015)



wirtschaftlichen Stärken Deutschlands auf dem Gebiet der Automatisierungstechnik, sie verdankt sich aber auch einer vielfältigen Forschungslandschaft, in der nicht zuletzt wesentliche Beiträge zur Entwicklung autonomer Roboter, KI und ML gemacht worden sind. Als herausragende deutsche Zentren für Forschung und Entwicklung sind diesbezüglich zu nennen (dazu und zum Folgenden Behnke 2014, S. 16 ff.):

- Das DLR Robotik und Mechatronik Zentrum (RMC),¹¹¹ das Forschung und Entwicklung auf den Gebieten Robotik, Mechatronik und optische Systeme für Anwendungen in den Bereichen Raumfahrt, Luftfahrt und Verkehr betreibt. Hier wurden beispielsweise die Leichtbauroboter entwickelt, die inzwischen von der Firma Kuka vermarktet werden. Weitere herausragende Entwicklungen sind der Serviceroboter Justin, der medizinische Roboter MIRO und die Laufmaschine TORO.
- Das Max-Planck-Institut für Intelligente Systeme,¹¹² das an den Standorten Tübingen und Stuttgart grundlegende Arbeiten mit dem Ziel des Verständnisses der Prinzipien von Wahrnehmen, Lernen und Handeln in autonomen Systemen durchführt. Das Institut verfolgt einen interdisziplinären Ansatz. Besonders relevant für die hier besprochene Thematik sind die Arbeiten der Abteilungen Perzeptive Systeme, Empirische Inferenz und Autonome Motorik.
- Auch die Fraunhofer-Gesellschaft betreibt angewandte Forschung auf dem Gebiet der Robotik und KI, unter anderem am Institut für Produktionstechnik und Automatisierung IPA,¹¹³ an dem z. B. die Care-O-Bot-Serviceroboter entwickelt wurden, oder am Institut für Optronik, Systemtechnik und Bildauswertung IOSB,¹¹⁴ wo Roboter zur Unterstützung von Rettungskräften entwickelt werden. Am Institut für Intelligente Analyse- und Informationssysteme IAIS werden die Forschungsgebiete Kognitive Robotik und Deep Learning verfolgt.¹¹⁵
- Das Deutsche Forschungszentrum für Künstliche Intelligenz (DFKI),¹¹⁶ das Standorte in Kaiserslautern, Saarbrücken und Bremen unterhält und auf dem Gebiet innovativer Softwaretechnologien auf der Basis von KI-Methoden die führende wirtschaftsnahe Forschungseinrichtung Deutschlands ist. Im Bremer Robotics Innovation Center (RIC) werden mobile Robotersysteme entwickelt, die an Land, zu Wasser, in der Luft und im Weltraum für komplexe Aufgaben eingesetzt werden.

111 www.dlr.de/rmc/rm (15.12.2015)

112 www.is.mpg.de/de (15.12.2015)

113 www.ipa.fraunhofer.de (15.12.2015)

114 www.iosb.fraunhofer.de/servlet/is/11/ (15.12.2015)

115 www.iais.fraunhofer.de (15.12.2015)

116 www.dfki.de (15.12.2015)



Daneben gibt es auch an vielen Universitäten führende Forschergruppen (dazu und zum Folgenden Behnke 2014, S. 16 f.). Hervorzuheben sind die TU München, wo bis 2014 im Rahmen des DFG-Exzellenzclusters Kognition für technische Systeme (CoTeSys)¹¹⁷ Serviceroboter für den Haushalt entwickelt wurden, sowie das Institut für Anthropomatik und Robotik¹¹⁸ des Karlsruher Instituts für Technologie (KIT), an dem im Rahmen des DFG-Sonderforschungsbereichs »Humanoide Roboter«¹¹⁹ (SFB 588, Laufzeit 2001–2012) die Armar-Serviceroboter entstanden sind. Eine umfangreiche Liste der zahlreichen deutschen Arbeitsgruppen findet sich im Internet.¹²⁰

Als besondere Stärke der deutschen Robotikforschung gilt ihre breite akademische Basis, in der grundlagenorientiert an autonomen Robotersystemen gearbeitet wird (dazu und zum Folgenden Behnke 2014, S. 18 f., sowie Schaal 2014, S. 75). Im Hinblick auf die autonome Robotik wird allerdings als Manko gesehen, dass im Bereich der Industrie noch die klassische Automatisierung im Vordergrund steht – die Industrieforschung in der autonomen Robotik ist in anderen führenden Robotikländern deutlich stärker, beispielsweise in den USA (z. B. Google, iRobot, Rethink Robotics) oder in Japan (z. B. Yaskawa, Kawada, Fanuc, Toyota, Honda). Hinzu kommt, dass sich in Deutschland zwischen akademischer Forschung und der Industrie eine recht große Lücke auftut, welche eine Überführung der Forschungsergebnisse in Innovationen erschwert – die intensive Kooperation zwischen dem DLR und Kuka, die u. a. den Leichtbauroboter hervorgebracht hat, ist hier eine der wenigen erwähnenswerten Ausnahmen (Bischoff et al. 2010). Um den tiefen Graben zwischen Grundlagenforschung und Anwendung zu überwinden, bräuchte es eine gezielte und risikobereite Förderung vielversprechender Forschungsprojekte der autonomen Robotik in einer finanziell signifikanten Größenordnung – die fehlt aber in Deutschland weitgehend. Zwar gab es von 2005 bis 2009 das Verbundprojekt »Deutsche Servicerobotik Initiative« (DESIRE),¹²¹ das im Rahmen des Programms »Leitinnovation Servicerobotik« des Bundesministeriums für Bildung und Forschung (BMBF) gefördert wurde. Auch der aktuelle BMBF-Förderschwerpunkt »Mensch-Technik-Interaktion im demografischen Wandel« weist vielfältige inhaltliche Bezüge zur Servicerobotik auf.¹²² Dennoch werden vom BMBF derzeit aber nur sieben kleinere, anwendungsorientierte Servicerobotikprojekte geför-

117 www.cotesys.org/ (15.12.2015)

118 www.informatik.kit.edu/1323.php (15.12.2015)

119 <http://gepris.dfg.de/gepris/projekt/5484926> (15.12.2015)

120 www.robotergesetze.com/links/ (15.12.2015)

121 www.service-robotik-initiative.de/ (15.12.2015)

122 Die Ausrichtung der deutschen Forschungsförderung im Bereich Robotik wird vertiefend in dem Horizon-Scanning des VDI/VDE-IT zum Thema »Mensch-Maschine-Entgrenzung im Anwendungsfeld ›Gesundheit im demografischen Wandel« betrachtet (in Vorbereitung).



dert,¹²³ was deutlich hinter den Programmen anderer Länder zurückbleibt. So hat die USA neben der DARPA-Förderung eine »National Robotics Initiative« ins Leben gerufen, in der autonome Roboter entwickelt werden sollen, die Schweiz hat mit dem National Center of Competence Research (NCCR) ein Robotikkompetenzzentrum gegründet und auch Japan hat kürzlich eine Robotikstrategie vorgelegt.

Fazit

2.5

Die autonome Robotik hat in den letzten Jahren zweifelsohne erhebliche Fortschritte gemacht, was sich an der wachsenden Verbreitung autonomer Systeme in Forschung, Industrie und Haushalt ablesen lässt. Die Breite der dargestellten Anwendungsbereiche macht zudem deutlich, dass das Feld das Potenzial hat, viele alltägliche Lebensbereiche grundlegend zu verändern (Behnke 2014, S.37). Auch wenn autonome Roboter heute schon über erstaunliche Fähigkeiten in einzelnen Aspekten verfügen, ist dennoch festzuhalten, dass noch starke Beschränkungen bestehen, die auf längere Sicht Gegenstand intensiver Forschungen bleiben dürften. So beschränkt sich der kommerzielle Aufschwung im Wesentlichen auf mobile Roboter, d.h. Systeme, die sich auf Rädern fortbewegen und nur sehr spezialisierte Fertigkeiten haben (dazu und zum Folgenden Schaal 2014, S. vi f.). Die Königsdisziplin der autonomen Robotik, nämlich humanoide Systeme, die auch komplexere Manipulationsaufgaben in unstrukturierten Umgebungen selbstständig durchführen können, befindet sich nach wie vor in einem frühen Forschungsstadium – und damit auch Assistenz- und Serviceroboter, die über diese Kompetenzen verfügen müssen, um die in sie gesetzten großen Zukunftshoffnungen erfüllen zu können. Die Forschungsaufgaben, die es diesbezüglich noch zu lösen gilt, sind groß und sind auf mindestens zwei unterschiedlichen Ebenen angesiedelt (zum Folgenden Schaal 2014, S.72 f.):

- › Die Verbesserung der Hardware wie effizientere Rechner, längerfristig verfügbare Energiequellen, beweglichere und robustere Aktuatoren sowie eine leistungsstärkere Sensorik: Die bestehenden Hardwarekomponenten sind oft noch zu groß und teuer sowie nicht auf den erforderlichen Leichtbau ausgelegt, und insbesondere die mechatronische Integration der unzähligen Bausteine zu einem ebenso kompakten wie funktionalen Gesamtsystem stellt eine große Hürde dar.
- › Die Entwicklung standardisierter Softwarearchitekturen, welche eine effiziente Programmierung von Wahrnehmungs-, Lern- und Anpassungsfähig-

¹²³ www.softwaresysteme.pt-dlr.de/de/forschungsvorhaben-servicerobotik.php
(15.12.2015)



keiten ermöglicht: Diesbezüglich ist oft noch nicht einmal geklärt, wie robuste Perception-Action-Learning-Systeme überhaupt zu programmieren sind, d.h. welche Algorithmen sich hierfür eignen.

Die Entwicklung autonomer Robotiksysteme erfordert demzufolge einen besonders langen Atem und geeignete institutionelle sowie finanzielle Rahmenbedingungen. Viele Einzelprojekte scheitern mangels langfristiger Finanzierung oder weil sie es nicht schaffen, die benötigten interdisziplinären Fachkompetenzen zu bündeln, die für die Bewältigung eines solchen Vorhabens unabdingbar sind. Die weltweit hervorstechenden Forschungsprojekte in der autonomen Robotik zeichnen sich dadurch aus, dass eine beachtliche finanzielle Projektförderung führende Forschungsteams längerfristig zusammenbringt – beispielhaft hierfür ist die DARPA-Förderung des letzten Jahrzehnts, welche die USA in wenigen Jahren an die Spitze der Robotikforschung katapultiert hat (dazu und zum Folgenden Schaal 2014, S. vii u. 61 f.). Viele Staaten – insbesondere die USA und Japan – haben dies erkannt und entsprechende nationale Förderprogramme ins Leben gerufen. Auch die Privatindustrie setzt bereits seit Längerem auf das ökonomische Potenzial der Robotik und unterstützt langfristige und intensive Forschungsbemühungen. Zu beobachten ist, dass sich der Fokus der Unternehmen zunehmend von der klassischen Industrieautomatisierung auf die autonome Robotik mit ihren flexibleren Systemen verschiebt. Neben den namhaften Automobilherstellern und großen Zulieferern, die alle an Fahrerassistenzsystemen arbeiten, ist es in Deutschland vor allem die in Augsburg ansässige Firma Kuka, die sich maßgeblich in der Servicerobotik engagiert (Behnke 2014, S. 18). Auffällig ist jedoch, dass in den letzten Jahren vor allem die großen amerikanischen IT-Konzerne hohe Investitionen in die autonome Robotik getätigt haben. So hat Google mehrere kleine Robotikfirmen aufgekauft – darunter Boston Dynamics, welche die Roboter »Little Dog« und »Atlas« herstellt – und eine neue Robotiksparte aufgebaut. Amazon hat die Firma Kiva Robotics übernommen, um die Automatisierung seiner Warenlager weiterzutreiben. Außerdem arbeitet der Konzern aktiv an semiautONOMEN Drohnen sowie an der Verbesserung autonomer Manipulationsroboter.

Da die entsprechenden Forschungsaktivitäten und Entwicklungsbemühungen der Internetkonzerne teilweise vertraulich ablaufen (Schaal 2014, S. 62), gibt es Raum für kritische Spekulationen bezüglich deren wirtschaftlicher Motive und Zielsetzungen. Eine immer wieder geäußerte Befürchtung ist, dass sich die gewaltige Wirtschaftskraft dieser Konzerne in den Dienst von technologischen Machbarkeitsvisionen stellt, die demokratisch nicht legitimiert sind (Wagner 2015). Dass derartige Spekulationen nicht ganz unbegründet sind, zeigt sich daran, dass Google vor einigen Jahren den Transhumanisten Ray Kurzweil als technischen Direktor verpflichtet hat. Transhumanistische Zukunftsvisionen, nach denen autonome Systeme bald ein Selbstbewusstsein entwickeln und sich



bewusst der menschlichen Kontrolle entziehen könnten, erweisen sich bei nüchternen Betrachtung der Sachlage aber derzeit als kaum realistisch. Die jüngeren, ohne Zweifel beeindruckenden Anwendungserfolge, die auf Methoden der KI-Forschung basieren, beschränken sich im Wesentlichen auf körperlose Softwareagenten – Googles erfolgreiche Suchmaschine ist dafür sicherlich das beste Beispiel. Doch sind derartige KI-Systeme tatsächlich autonom, geschweige denn intelligent zu nennen? Angesichts ihrer rein virtuellen Existenz und ihrer sehr eng begrenzten »kognitiven« Leistungsfähigkeit ist dies äußerst fraglich. Angesichts dessen scheint absehbar weniger die Machtübernahme intelligenter Maschinen das drängende Problem zu sein als vielmehr die Tatsache, dass zwar immer einflussreichere, aber verhältnismäßig »dumme« Algorithmen zur Geschäftsbasis neuer digitaler Monopole werden und damit zu einem Machtwerkzeug, das sich gesellschaftlicher Kontrolle weitgehend entzieht (Lanier 2014).

Mit Blick auf autonome Robotiksysteme bleibt festzuhalten, dass sie sich – anders als Softwareagenten – in der physischen Realität zu bewähren haben, was die zuvor beschriebenen, großen FuE-Herausforderungen mit sich bringt. Vor diesem Hintergrund ist nicht davon auszugehen, dass sich die gesellschaftlichen Veränderungen abrupt vollziehen werden; eher ist mit einer allmählichen Verbreitung autonomer Roboter im Alltag zu rechnen, womit Raum für politische Weichenstellungen und eine gesellschaftliche Gestaltung der Entwicklung bleiben dürfte (Schaal 2014, S.70). Die gesellschaftliche Handhabe ist schon allein deshalb gegeben, weil sich der derzeitige Aufschwung der autonomen Robotik nicht in einem rechtsfreien Raum abspielt, sondern durch zahlreiche nationale und internationale Regelwerke eingehegt wird (dazu und zum Folgenden Behnke 2014, S.32 ff.). So haben autonome Roboter, die in engeren Kontakt mit Menschen kommen könnten, nach derzeitiger Rechtslage bestimmten Sicherheitsvorschriften und -normen zu genügen. Die hier maßgebliche Rechtsbestimmung ist die EU-Maschinenrichtlinie 2006/42/EG, die grundlegende Sicherheitsanforderungen für die Konstruktion und den Bau von Maschinen festlegt. Die Zulassung kann demzufolge nur erteilt werden, wenn eine Risikobeurteilung durchgeführt und die sich daraus ergebenden notwendigen Sicherheitsmaßnahmen umgesetzt wurden (VDMA 2014) – diese Kriterien stellen für autonome Systeme im Prinzip hohe Hürden dar. Derzeit ist jedoch noch alleine der Hersteller für die Sicherheit der Systeme verantwortlich (sogenannte Eigenzertifizierung), generell mangelt es noch an professionellen Stellen, welche die komplexe Zertifizierung von Robotersystemen unterstützen oder gar übernehmen könnten.

Selbstverständlich sind bestehende Vorschriften und Normen, die ja nach Anwendungsfeld stark differieren können, nicht in Stein gemeißelt, sondern müssen an sich verändernde Problemlagen angepasst werden. Mit dem Eindringen autonomer Roboter in alltägliche Lebensbereiche stellen sich hier viele



neue Fragen, nicht nur im Hinblick auf Sicherheitsaspekte, sondern auch bezüglich Haftungs- und Datenschutzfragen (Müller 2014). Allerdings gilt es im Auge zu behalten, dass der Grat zwischen einer Unterregulierung einerseits, die rechtliche Grauzonen schafft und Sicherheitsbedürfnisse ignoriert, sowie einer restriktiven Überregulierung andererseits, durch welche die Innovationstätigkeit in bestimmten Anwendungsbereichen ganz zum Erliegen kommen kann, unter Umständen sehr schmal ist (Royakkers/van Est 2015, S.332). Dieses Spannungsverhältnis, das auch als Collingridge-Dilemma¹²⁴ bekannt ist, ist typisch für die Regulation von emergierenden Technologien wie der autonomen Robotik. Um in solchen Situationen eine demokratische Governance von Innovationsprozessen zu ermöglichen, hat die Europäische Kommission vorgeschlagen, den Fokus weg von der Gestaltung der regulativen Rahmenbedingungen und hin zur Gestaltung des Forschungs- und Innovationsprozesses selber zu verschieben (von Schomberg 2014). Ziel dieser sogenannten »Responsible Research and Innovation« (RRI) ist u. a., die Bevölkerung möglichst früh in die Technologieentwicklung einzubeziehen, um Ergebnisse zu erhalten, die in Einklang mit fundamentalen Werten und gesellschaftlichen Bedürfnissen stehen. Dies geschieht in Anerkennung der Tatsache, dass Technologien immer, wenn auch oft nur implizit, bestimmte Wert- und Normvorstellungen verkörpern, insofern sie durch ihr spezifisches Design und ihre spezifische Funktionalität konkrete Nutzungsweisen nahelegen oder gar vorschreiben (und andere hingegen ausschließen); autonome Roboter, die sensible soziale Bereiche (teil)automatisieren (und damit zwangsläufig auch normieren) sollen, sind hierfür sicherlich ein besonders gutes Beispiel (van den Berg 2011). Es geht also darum, diese oft verborgenen Werte und Normen gesellschaftlich zu öffnen, indem die ethischen und gesellschaftlichen Dimensionen der neuen Technologien und ihrer Anwendungen möglichst frühzeitig reflektiert werden. In Bezug auf die autonome Robotik scheint dieser Reflexionsprozess besonders drängend. Denn Ziel dieses FuE-Feldes ist es ja, Maschinen zu erzeugen, die bisher dem Menschen vorbehaltene Tätigkeiten übernehmen. Diese menschenähnlichen Maschinen transportieren (im Rahmen technologischer Möglichkeiten) ein bestimmtes Bild des Menschen sowie sozialer Beziehungen und werden somit, je mehr sie in unseren Alltag Einzug halten, auch unser Selbstverständnis grundlegend herausfordern – in Bezug auf so zentrale Konzepte wie Selbstbestimmung, Verantwortungsfähigkeit und Identität.

124 Der britische Technikforscher David Collingridge (1982) machte als erster darauf aufmerksam, dass die Abschätzung von Technikfolgen einem grundsätzlichen Dilemma unterliegt: So sind die gesellschaftlichen Auswirkungen einer Technologie in einem frühen Stadium kaum absehbar – sie sind es vielmehr erst dann, wenn die Technologieentwicklung bereits weit fortgeschritten und damit kaum noch steuerbar ist.



Tendenzen zur Verschmelzung von Mensch und Maschine, die Thema des vorliegenden Berichts sind, werden derzeit hauptsächlich durch Entwicklungen in zwei hochdynamischen Technologiefeldern verstärkt:

- > Auf der einen Seite durch das weite Feld der Neurotechnologien, das die direkte Kopplung des menschlichen Gehirns bzw. Nervensystems mit maschinellen Systemen unterschiedlicher Art zum Ziel hat. Gegenwärtig geschieht dies noch vorwiegend aus therapeutischen und diagnostischen Gründen, teilweise bereits mit beachtlichem Erfolg. Die Brisanz neurotechnologischer Entwicklungen steht außer Frage, denn mit dem Gehirn wird nicht mehr nur der menschliche Körper, sondern im Prinzip das menschliche Selbst dem technischen Gestaltungs- und Optimierungswillen unterworfen.
- > Auf der anderen Seite arbeiten die Robotik- und die KI-Forschung an der Entwicklung von Maschinen, die möglichst autonom – d. h. quasi intelligent und folglich menschenähnlich – agieren. Insbesondere humanoide Manipulationsroboter, die zukünftig vermehrt auch in Alltagssituationen zum Einsatz kommen sollen, wecken Assoziationen zum künstlichen Menschen und verwischen, zumindest perspektivisch, zunehmend die Grenzen zwischen Maschine und Mensch. Auch wenn von einem breiteren Einsatz derartiger »Maschinenmenschen« noch nicht zu sprechen ist, steht außer Frage, dass die autonome Robotik unsere Gesellschaft zukünftig tiefgreifend prägen wird.

Noch entwickeln sich die beiden emergierenden Felder weitgehend abgegrenzt voneinander, was aufgrund der sehr unterschiedlichen wissenschaftlich-technologischen Grundlagen auch nicht weiter erstaunlich ist. Aus übergeordneter Sicht ist mit Blick auf ihre Implikationen jedoch hervorzuheben, dass ihre Anwendungsperspektiven die beiden Pole einer übergreifenden Entgrenzungsdynamik markieren. Mit anderen Worten, Cyborg und humanoider Roboter sind spiegelbildliche Reflexionsfiguren, die auf die reziproke Annäherung von Mensch und Maschine verweisen (Takahashi 2010): Maschinen werden immer menschenähnlicher (Roboter), gleichzeitig werden Menschen zunehmend technisiert (Cyborg).

Auf die Spitze getrieben wird diese Entwicklung durch die sich immer deutlicher abzeichnende Konvergenz von Robotik bzw. KI und Neurotechnologien. So sind Prothesen, welche die Bewegungsabsichten des Trägers aus seinen Muskelsignalen grob »herauslesen«, bereits Standard. Neuerdings wird aber auch intensiv an sogenannten intelligenten Roboterprothesen geforscht, die über



einen wesentlich größeren Grad an Autonomie verfügen, um auch bei schwachen Signalen die volle Funktionsfähigkeit von Hand, Arm oder Bein wiederherzustellen. Demzufolge gibt der Mensch nur den Bewegungsimpuls, den anschließenden Bewegungsablauf (z. B. Greifprozess) vollzieht die Prothese weitgehend selbstständig mithilfe vielfältiger Sensorik und maschineller Lernalgorithmen (Schaal 2014, S. 22 f.). Ähnliches gilt für weitere Anwendungen, die sich im Schnittfeld von Robotik und Neurotechnologien herauszukristallisieren beginnen,

- › etwa intelligente Implantate (Elektrozeutika), die Körperfunktionen überwachen und therapeutische Maßnahmen nach Bedarf autonom einleiten (Kap. III.1.4),
- › robotische Exoskelette (wie sie bereits in der Rehabilitationstherapie eingesetzt werden; Sale et al. 2012), welche die Laufbewegung und Balance selbstständig regeln, oder
- › humanoide Roboter, die mittels BCI gesteuert werden – bislang allerdings nur im Labor (Gopura et al. 2011; Petit et al. 2015).

Das Ergebnis ist jeweils, dass Mensch und autonom agierende Technik zu einer zunehmend schwerer auflösbaren Einheit verschmelzen, sodass in letzter Konsequenz nicht mehr eindeutig unterscheidbar ist, ob der Mensch die Maschine steuert oder umgekehrt. Während diese Anwendungsfelder gute Beispiele dafür sind, wie sich der Mensch die Technik gewissermaßen »einverleibt«, ist gleichzeitig zu beobachten, dass sich die Entwicklung autonomer Systeme immer stärker an menschlichen Prinzipien der Reiz- und Informationsverarbeitung, des Lernens oder der Aufmerksamkeitssteuerung orientiert (Gransche/Warnke 2014, S. 6) – Beispiele sind etwa das »Neuromorphic Computing«¹²⁵ (Waldrop 2013), wie es im »Human Brain Project« schwerpunktmäßig erforscht wird.¹²⁶

Zukunftsvisionen und ihre gesellschaftliche Relevanz

Das visionäre Potenzial dieser Entwicklungen steht außer Frage. Angesichts der rasant wachsenden Rechenkapazität von Computern sowie der zunehmenden Miniaturisierung relevanter elektrotechnischer Komponenten werden Zukunftserwartungen immer populärer, welche im Kern die baldige und weitgehende Entgrenzung von Mensch und Maschine prognostizieren. In ihrer extremen Form, wie sie zum Beispiel von Ray Kurzweil und anderen Transhumanis-

¹²⁵ Mit neuromorphen Computersystemen wird versucht, die plastische Struktur des Gehirns, insbesondere dessen neuronale Verschaltungsarchitektur, auf Silizium nachzubauen. Die Hoffnung ist vor allem, dass neuromorphe Chips deutlich energieeffizienter arbeiten als bislang verfügbare Systeme.

¹²⁶ www.humanbrainproject.eu/de/neuromorphic-computing-platform (15.12.2015)



ten seit längerem vertreten wird, prophezeien diese Visionen den Anbruch eines neuen posthumanistischen Zeitalters, in dem die menschliche Natur durch die gänzliche Verschmelzung von Mensch (bzw. Gehirn) und superintelligenten Maschinen technisch überwunden bzw. völlig neu definiert wird. Derartige, zuweilen quasireligiös anmutende Visionen bilden wesentliche Referenzpunkte der gesellschaftlichen, aber auch der ethischen Debatte rund um Entwicklungen im Bereich der Neurotechnologien, der Robotik und der KI. Diese Debatte folgt damit älteren Technikdebatten wie z. B. der über die Nanotechnologie, mit denen sie zudem gemein hat, dass die in ihr zirkulierenden Visionen sich im Wesentlichen an ältere Schriften (der 1920er bis 1960er Jahre) anlehnen und lediglich technisch z. T. aktualisiert worden sind.

Dabei differieren die Auffassungen über die gesellschaftlichen Chancen und Risiken mitunter erheblich, was sich gerade aktuell am Beispiel des öffentlichen Diskurses um die Robotik deutlich nachverfolgen lässt. Große Hoffnungen auf eine neue industrielle Revolution und vielseitige Erleichterungen im Alltag treffen auf düstere Zukunftsbilder, die eine Massenarbeitslosigkeit aufgrund fortschreitender Automatisierung oder eine nicht mehr beherrschbare »Explosion« maschineller Intelligenz prognostizieren. Dabei zeigt sich bei genauerer Betrachtung nicht selten, dass sich konkrete Zukunftsprognosen auf Einschätzungen der Entwicklungsdynamik stützen, die bestenfalls hochkontrovers, manchmal auch unklar sind oder gar jeglicher Grundlage entbehren. Einer der häufigsten »futuologischen Fehlschlüsse« (Sturma 2003, S. 38) besteht darin, den technischen Entwicklungsprozessen eine »deterministische Eigenlogik« zu unterstellen (Grunwald 2011, S. 6) – etwa, indem die atemberaubende Geschwindigkeit des digitalen technischen Fortschritts extrapoliert und davon auf die baldige Überwindung technologischer Hürden geschlossen wird. Daraus, dass die Rechenleistung von Prozessoren exponentiell anwächst, folgt allerdings nicht, dass wissenschaftlich-technische Fragen – etwa die Entschlüsselung des Gehirns – auch tatsächlich durch pure Rechenkraft alleine lösbar sind (Lanier 2014). Zudem wird gerne übersehen, dass Technikentwicklung auch ein sozialer Prozess ist, geprägt von vielfältigen gesellschaftlichen und ökonomischen Faktoren, wie die oft schwierige Übertragung von Laborlösungen in die praxistaugliche Anwendung deutlich zeigt. Dennoch neigen relevante Akteure angesichts erster Erfolge immer wieder auch zu kalkulierten Übertreibungen, um die aufmerksamkeitsgenerierenden Mechanismen der Mediengesellschaft für ihre Sache zu nutzen. So entsteht eine diskursive Gemengelage, in der sich »Facts and Fiction« oft kaum noch unterscheiden lassen, was insgesamt nicht zu einem verantwortungsvollen Umgang mit dem technischen Fortschritt und seinen gesellschaftlichen Herausforderungen beiträgt, sondern die Unsicherheiten und Unklarheiten tendenziell erhöht.



Überschießende Erwartungen sind nicht zuletzt ein Symptom der noch sehr unscharfen Konturen der technologischen Felder selbst, die hinsichtlich möglicher Anwendungsperspektiven einen weiten Zukunftshorizont eröffnen und damit allerlei Spekulationen Nahrung bieten. Daraus folgt umgekehrt, dass jegliche Prognosen zu Technikentwicklungen und mehr noch eine Antizipation ihrer Folgen auf äußerst schwankendem Grund stehen – dies zeigen auch die Erfahrungen aus ähnlich gelagerten Debatten zu anderen emergierenden Technologiefeldern, etwa der Nanotechnologie. Eine Schlussfolgerung daraus lautet, dass eine Zukunftsorientierung praktisch nur über eine reflexive Betrachtung gegenwärtiger Entwicklungen in Wissenschaft und Gesellschaft erreicht werden kann: zum einen, indem man sich mit »gesellschaftlichen Praktiken, unterschwelligen Sorgen, impliziten Hoffnungen und Befürchtungen« beschäftigt (Grunwald 2015, S.66), wie sie in den erwähnten Technikvisionen zum Ausdruck kommen. Da diese, so spekulativ sie auch sind, die weitere Entwicklungsdynamik subtil beeinflussen können (auch dies eine Folge der sozialen Bedingtheit von wissenschaftlich-technischen Entwicklungsprozessen), sind sie nämlich gleichwohl von faktischer Bedeutung. Zum anderen ist es wichtig, und darin besteht eine wichtige Zielsetzung dieses Berichts, eine kritische Bestandsaufnahme der aktuellen Leistungsfähigkeit der Technologien einerseits und ihrer sich bereits abzeichnenden Anwendungsmöglichkeiten andererseits vorzunehmen. Nur vor diesem Hintergrund lässt sich jenseits spekulativer Erwägungen eine fundierte Einschätzung gewinnen, welche gesellschaftspolitischen Weichenstellungen bereits jetzt geboten scheinen und wie ein verantwortungsvoller Umgang mit dem wissenschaftlich-technischen Fortschritt aussehen könnte.

Vor diesem Hintergrund ist festzuhalten, dass nach derzeitigem wissenschaftlich-technischem Entwicklungsstand viele der in den Zukunftsdiskursen formulierten Erwartungen weitgehend spekulativ erscheinen. Zwar erlauben es die bestehenden Neurotechnologien schon jetzt, die menschlichen Sinnesfähigkeiten zu modifizieren oder zu ergänzen. Aktuelle therapeutische wie spielerische Anwendungen solcher Cyborgtechnologien lassen es aber nicht als angebracht erscheinen, hier von einer wünschbaren Steigerung der menschlichen Leistungsfähigkeit oder gar einer Verbesserung der menschlichen Natur zu sprechen. In erster Linie deshalb, weil die Zusatzfähigkeiten (wie z. B. die Wahrnehmung von Ultraviolett mithilfe der Eyeborg-Technologie Neil Harbisons) z. T. gesellschaftlich und sogar für den individuellen Nutzer kaum von relevantem Nutzen und darüber hinaus (insbesondere bei den invasiven Technologien) damit gesundheitliche Risiken verbunden sind, die heute noch keine Anwendung bei Gesunden rechtfertigen und sicherlich auch in den nächsten 10 Jahren nicht erwarten lassen. Überdies erscheint es problematisch, dass dezidiert transhumanistische Autoren in ihren Zukunftsvisionen nicht nur einer Steigerungs- und Entleiblichungslogik folgen, sondern oft auch die von ihnen



imaginierten Zukünfte als kaum aufhaltbar darstellen. Hier stellt man eine Einseitigkeit zukunftsbezogenen Denkens fest, die nicht nur bedenklich ist, weil viele dieser Visionen kaum auf dem gegenwärtigen Forschungs- und Entwicklungsstand basieren, sondern vor allem auch, weil dadurch reale Chancen und Herausforderungen durch die Technikentwicklung aus dem Blickfeld zu geraten drohen.

Was die KI und Robotik anbelangen, so dürfte die Aussicht auf eine Machtübernahme künstlicher Intelligenzen derzeit vernachlässigbar sein angesichts der großen technologischen Herausforderungen, vor denen z. B. die Entwicklung komplexerer Manipulationsroboter steht. Bei allem Respekt für die intellektuelle Statur derjenigen, die – mit welcher Risiko-Nutzen-Abwägung auch immer – die Zukunftsvision einer sehr »starken« KI zu einem wichtigen, wenn nicht zentralen Thema heutiger politisch-gesellschaftlicher und akademischer Reflektion machen wollen, ist doch große Skepsis hinsichtlich ihrer Bedeutung für die Gegenwart oder nähere Zukunft angebracht. Vor allem ist weitgehend unklar, wie sich eine »starke« KI realisieren ließe. Die bisherigen Erfolge, welche die KI-Forschung unzweifelhaft vorzuweisen hat, beschränken sich auf lernfähige Softwareanwendungen (etwa zur Sprach- und Bilderkennung, als Übersetzungs- oder Spielprogramme, im Rahmen des automatisierten Börsenhandels etc.; Bostrom 2014, S.27 ff.), die zwar das digitale Zeitalter tiefgreifend prägen, aber aufgrund ihrer Spezialisierung gleichwohl vollends dem Paradigma der »schwachen« KI verhaftet sind. Gängige Übersetzungsprogramme beispielsweise basieren auf rein statistischen Verfahren der Textanalyse. Weder verstehen sie die Sprachen, die sie übersetzen, noch sind sie sich überhaupt bewusst, dass sie übersetzen. Auch wenn es sich hierbei inzwischen um mächtige Werkzeuge handelt – intelligentes oder autonomes Verhalten im menschlichen Sinne zeigen sie in keiner Weise. Dies ist nicht zuletzt deshalb so, weil ihre Funktionalität rein auf die digitale Datenwelt beschränkt bleibt. Die Herausbildung von Selbstbewusstsein und (allgemeiner) Intelligenz könnte jedoch wesentlich an die Möglichkeit gebunden sein, so legen es zumindest neuere Befunde aus den Kognitionswissenschaften nahe, körperlich mit der physischen (und sozialen) Umwelt zu interagieren – genau die dafür erforderliche funktionale Integration von Wahrnehmungs-, Planungs- und Manipulationsfähigkeiten gehört aber zu den größten und bislang weitgehend ungelösten Problemen der KI- und Robotikforschung. Mit dem kommerziellen Durchbruch von Servicerobotern, die sich auch in alltäglichen Situationen intelligent verhalten und selbstständig zurechtfinden, ist in den nächsten 20 Jahren, und wahrscheinlich noch darüber hinaus, jedenfalls nicht zu rechnen.

Allein der – durchaus bemerkenswerte – Umstand, dass mit global führenden Unternehmen und Holdings wie Alphabet (vormals Google) einige Akteure in diesem Diskurs stark präsent sind, denen am ehesten eine Realisierung der



Visionen zugetraut werden kann, reicht als Begründung für eine antizipatorische Governance noch nicht aus. Es ist vielmehr hier – wie auch bei den neurotechnologischen Enhancementdiskussionen – darauf hinzuweisen, dass die eher an weit in die Zukunft blickende Science-Fiction gemahnenden Visionen, die ja durchaus zu faszinieren vermögen, dazu beitragen können, realistischere Chancen und Risiken aus dem Blickfeld zu verdrängen. Die zu beobachtende Verkürzung der öffentlichen Debatte auf Cyborgs und künstliche Superintelligenzen, die außer Kontrolle geraten könnten, ist vor allem deshalb wenig hilfreich, weil sie den Blick von den eigentlich drängenden normativen Fragen ablenkt, die mit der weiteren Verbreitung autonomer Systeme bereits jetzt einhergehen.

Anthropologische Grenzerfahrungen und ihre normativen Implikationen

Dafür, dass in absehbarer Zeit die technische Optimierung der »Natur« des Menschen (in einem größeren, d.h. gesellschaftlich relevanten Maßstab) oder eine »Intelligenzexplosion« bei Maschinen bevorsteht, gibt es demnach keine Anzeichen. Mithin stellt sich die Frage: Welcher Art sind denn die gesellschaftlichen Herausforderungen, die durch den schleichenden Trend zur technologischen Entgrenzung von Mensch und Maschine entstehen?

Aus Sicht des niederländischen Rathenau Instituut sind autonome Roboter und neurotechnologische Anwendungen paradigmatische Beispiele einer übergreifenden Entwicklung, die als »Intimate-technological Revolution« charakterisiert wird (van Est 2014). Deren Kernmerkmal bestehe darin, dass digitale Geräte – etwa Wearables, Kommunikationsplattformen, Neuroprothesen, Serviceroboter etc. – unsere Körper und sozialen Praktiken immer stärker durchdringen und damit verändern, vor allem aber auch subtil beeinflussbar machen. Das Rathenau Instituut spielt damit primär auf die zunehmende Verschmelzung von persönlicher und digitaler Sphäre an, eine Entwicklung, wie sie für das Zeitalter der Digitalisierung charakteristisch ist. Das räumliche Vordringen elektronischer Geräte in höchstpersönliche Lebens- und Körperbereiche wirft vor allem ethisch-rechtliche Fragen hinsichtlich des Schutzes von Persönlichkeitsrechten auf (Datenschutz, Schutz der Privatsphäre und der Autonomie) – zweifelsohne sind dies auch wichtige Aspekte der hier behandelten Technologien. Nicht nur neurotechnologische Anwendungen, sondern auch autonome Roboter, die eng mit Menschen zusammenarbeiten, benötigen für ihre Steuerung zahlreiche Informationen, die unter Umständen bis in die privatesten Bereiche der Nutzer hineinreichen (Behnke 2014, S. 34). Der spezifische Blick auf die in diesem Bericht behandelten Technikphänomene rückt aber noch etwas anders gelagerte Folgendimensionen in den Vordergrund, die sich weniger aus der räumlichen



Verschmelzung von Mensch und Maschine ergeben, sondern vor allem mit anthropologischen Grenzüberschreitungen zusammenhängen.

Entwicklungen im Bereich der Neurotechnologien und autonomen Robotik führen nämlich aus genau entgegengesetzter Richtung dazu, dass Mensch und Maschine sich zunehmend kategorial *angleichen*. Das heißt, die Problematik liegt weniger darin, dass die Technik sich dem Menschen immer mehr annähert oder ihm visuell ähnlicher wird – vielmehr geht es um die subtile Verwischung der fundamentalen Charakteristika, die eine Person oder eine Maschine ausmachen. Eine zentrale Rolle dabei spielt der Begriff der Handlung, der nach gängigen ethisch-rechtlichen Auffassungen eng mit dem Mensch- bzw. Personensein verknüpft ist (Neuhäuser 2012). So gelten ausschließlich Personen als handlungsfähig, weil nur ihnen die Fähigkeit zugeschrieben wird, sich aus freien Stücken in ihrem Tun Ziele zu setzen und die Mittel auszuwählen, um diese zu erreichen. Maschinen hingegen, als Produkte des Menschen, sind nach herkömmlichem Verständnis reine technische Hilfsmittel, um menschliche Zwecke zu erfüllen. Ihnen wird folglich nicht die Fähigkeit zuerkannt, eigene Absichten zu verfolgen, weshalb ihr Verhalten als reiner mechanischer Ablauf zu qualifizieren ist. Diese diametrale Abgrenzung von Mensch und Maschine, die sich wesentlich an Aspekten der Handlungs- und Entscheidungsfähigkeit festmacht, bildet einen wesentlichen Pfeiler unserer normativen Ordnung, etwa in Bezug auf moralische Verantwortungszuschreibungen, die auch rechtlichen Vorstellungen (zivilrechtliche Haftung, strafrechtliche Schuldfähigkeit) zugrunde liegen (Hilgendorf 2012). Während Personen für ihre Taten selbstverständlich zur Verantwortung gezogen werden, hat das »Fehlverhalten« von Maschinen, seien die Konsequenzen auch noch so gravierend, keinerlei juristische Konsequenzen für die Maschine selber – es werden höchstens Hersteller oder Nutzer haftbar gemacht.

Der weitere Fortschritt bei Robotern und Neurotechnologien bringt diese Grundordnung immer mehr ins Wanken, insofern nämlich, als Technik im Zuge dessen ihren rein instrumentellen Charakter verliert und als zunehmend selbstständig agierender Akteur ihrerseits tief in das persönlich-individuelle wie auch das gesellschaftlich-kollektive Selbstbewusstsein einzugreifen beginnt. So machen neurotechnologische Zugriffe auf das menschliche Gehirn eben nicht bei körperlichen Funktionen halt, sondern können auch mehr oder weniger direkte Auswirkungen auf psychische Funktionen und Persönlichkeitsaspekte der Betroffenen haben. Bei der tiefen Hirnstimulation etwa besteht das Problem von Persönlichkeitsveränderungen, was im Extremfall dazu führen kann, dass die Authentizität der psychischen Verfassung oder gar Zurechnungsfähigkeit eines Patienten infragegestellt wird – was aber nicht ausschließt, dass dieser sich genau in der technisch ausgelösten Verfassung authentisch und gut fühlt (z. B. Mackenzie 2011; Pham et al. 2015).



Je stärker das menschliche Selbst in Technisierungsprozesse einbezogen wird, umso mehr löst sich auch die traditionelle Dichotomie zwischen menschlicher Autonomie und maschineller Zweckgebundenheit auf. Sie erscheint spätestens dann fraglich, wenn autonome technische Systeme einen aktiven Part in menschlichen Handlungsvollzügen spielen, d.h. mit dem Menschen zu einer hybriden Handlungseinheit verschmelzen – wie es bei intelligenten Prothesen in nicht allzu ferner Zukunft der Fall sein könnte. Die klare Trennung zwischen menschlichem Akteur und technischem Hilfsmittel ist dann kaum noch möglich, was fundamentale anthropologische Kategorien – wie Verantwortungsfähigkeit, Selbstbestimmtheit, Identität – herausfordert. Daraus ergeben sich ganz neue moralische und rechtliche Fragen, die in ähnlicher Weise auch mit der weiteren Verbreitung selbstlernender Roboter virulent werden. Da diese sich zunehmend eigenständig verhalten, wird bereits intensiv diskutiert, wer für Schäden, die von autonomen Robotern verursacht werden, verantwortlich ist (Beck 2012). Ist es der Hersteller, für den das spezifische Verhalten des autonomen Systems (aufgrund dessen lernender Weiterentwicklung) unter den gegebenen Umständen ja kaum vorhersehbar war? Oder ist es der Nutzer, der den Roboter zwar eingesetzt hat, dessen Verhalten aber gar nicht kontrollieren konnte? In dem Maße, in dem Maschinen sich autonomer verhalten und in soziale Praktiken integriert werden, könnten diese beiden Alternativen zunehmend unbefriedigend erscheinen. Zwar sind aktuelle Robotersysteme noch weit davon entfernt, eine eigene Handlungspersönlichkeit zu entwickeln. Schon jetzt können sich bei teilautonomen Maschinen aber rechtliche Unsicherheiten in Zurechnungsfragen ergeben (wie ein aktueller Fall zeigt),¹²⁷ die sich in Zukunft bei steigender maschineller Autonomie weiter zuspitzen dürften und bereits heute die Frage nach dem moralischen und rechtlichen Status von autonomen Robotern laut werden lassen (Behnke 2014, S. 36).

Während die rechtlichen Konsequenzen, welche die »Übertragung unseres normativen Grundvokabulars auf Maschinen« mit sich bringt, nach Einschätzung von Rechtsexperten größtenteils beherrschbar scheinen (Hilgendorf 2012, S. 132) – auch Tieren und juristischen Personen werden schließlich Rechte und teilweise Pflichten zugeschrieben –, sind die moralischen Implikationen dieser anthropologischen Grenzverschiebungen noch nicht wirklich absehbar. Klar ist, dass sie traditionelle Orientierungen infrage stellen und damit normative Unsicherheiten aufwerfen – etwa im Hinblick auf das leitende Menschenbild, den wünschbaren Grad an gesellschaftlicher Automatisierung (und damit auch Normierung) sowie die Gestaltung des Mensch-Technik-Verhältnisses –, die sich bereits jetzt abzeichnen. Verstärkt werden diese Unsicherheiten dadurch,

¹²⁷ www.welt.de/wirtschaft/article146407129/Wen-soll-das-autonome-Auto-lieber-ueberfahren.html (15.12.2015)



dass zentrale Zielvorstellungen in einem eigentümlichen Spannungsverhältnis zueinander zu stehen scheinen: Auf der einen Seite ist die zunehmende Automatisierung von Maschinen in vielen gesellschaftlichen Bereichen erwünscht und wird angestrebt (z. B. aus wirtschaftlichen oder medizinischen Gründen), gleichzeitig ist damit auf der anderen Seite ein menschlicher Autonomie- und Kontrollverlust verbunden, der menschliche Grundbedürfnisse nach Sicherheit und Selbstbestimmung zu verletzen droht und die Frage nach den Grenzen des Fortschritts bzw. der angemessenen Gestaltung der Systeme aufwirft. Vor diesem Hintergrund scheint eine gesellschaftliche Verständigung darüber, ob es bestimmte Lebensbereiche und Aufgaben gibt, die mit großer Vorsicht zu behandeln sind oder gar ganz von einer Automatisierung ausgenommen werden sollten, schon jetzt dringend erforderlich (Royakkers/van Est 2015, S.322). Als Konsens gilt, dass die menschliche Würde unveräußerlich ist und Technisierungsprozesse, die auf eine Instrumentalisierung und Verdinglichung des Menschen hinauslaufen – was dann der Fall ist, wenn sich die ursprüngliche Zweck-Mittel-Logik der Technikanwendung ins Gegenteil verkehrt (der Mensch also zum Mittel degradiert wird, das der Erfüllung maschineller Zwecke dient) –, diesbezüglich eine wesentliche moralische Grenze überschreiten, die nicht überschritten werden sollte. Aus diesem Instrumentalisierungsverbot lässt sich die allgemeine Richtschnur ableiten, dass die Entwicklung technischer Mittel nicht zum Selbstzweck werden, sondern immer im Dienste des menschlichen Wohlergehens als übergeordnetem und maßgeblichem Zweck stehen sollte (Royakkers/van Est 2015, S.334). Wie sich diese abstrakte Forderung im Hinblick auf Regulierungsfragen und die Gestaltung des technischen Fortschritts fassbar machen lässt, ist damit natürlich noch keineswegs geklärt.

Derzeit dreht sich die normative Debatte weitgehend um Sicherheitsfragen und wird hauptsächlich auf technischer Ebene geführt, etwa im Rahmen nicht demokratisch legitimierter designtechnischer Normungsprozesse der internationalen Organisation für Normung (ISO). Vorrangig technokratisch angelegte Lösungsansätze sind der Entwicklung aber nicht angemessen, da es sich bei komplexen Technologien wie autonomen Robotern oder Neurotechnologien um soziotechnische Systeme handelt – Systeme also, die tiefgreifende Auswirkungen auf die sozialen Praktiken und Strukturen haben, in die sie eingebettet werden, und deshalb nicht unabhängig von ihren Anwendungskontexten zu beurteilen sind. Beispiele hierfür sind die soziokulturellen Auswirkungen der zunehmenden Verbreitung von Cochlea-Implantaten auf die Gehörlosengemeinschaft oder die Art und Weise, wie bereits so einfache teilautonome Maschinen wie Staubsaugerroboter eine robotergerechte Umgestaltung der häuslichen Umgebung erforderlich machen (Sung et al. 2007). Da die soziokulturellen Veränderungen, die durch neue technische Systeme angestoßen werden (neue Infrastrukturen, Standards, Handlungsrouinen, Kompetenzgewinne wie auch



-verluste) oft tiefgreifend und irreversibel sind – die Art und Weise, wie Smartphones in wenigen Jahren unsere Alltagskultur umgestaltet haben, ist hierfür sicherlich das eindrucklichste Beispiel –, sollte die normative Debatte öffentlich geführt werden und nicht nur die Technologien im engeren Sinne, sondern auch ihre soziale und kulturelle Anwendungsdimension thematisieren (Robo-Law 2014, S.209). Angesichts der anthropologischen Herausforderungen der neuen Technologien sollte es dabei letztlich immer auch um die Frage gehen, was wir als genuinen Wesenszug des Menschen und seiner Kultur verstehen, was davon wir für die Zukunft erhalten und vor maschinellm Zugriff schützen wollen.

Ausblick

Eine frühzeitige Auseinandersetzung mit den gesellschaftlichen Konsequenzen, aber auch Rahmenbedingungen der Entgrenzungsdynamik scheint angesichts ihrer großen normativen Herausforderungen dringend erforderlich. Wie eine derartige »antizipatorische Governance«, die nicht nur Wissenschaft und Technik, sondern auch Politik, Wirtschaft und potenzielle Nutzer einbeziehen sollte, organisiert werden könnte, ist noch eine weitgehend offene Frage. Aus der Tatsache, dass die normativen Implikationen der neuen Entgrenzungstechnologien nur kontextbezogen zu beurteilen sind, lässt sich zumindest folgern, dass eine übergeordnete Debatte zu den Chancen und Risiken dieser Entwicklungen wenig sinnvoll erscheint. Stattdessen ist die gesellschaftliche Auseinandersetzung sinnvollerweise mit einem differenzierten Blick auf die jeweiligen Anwendungsfelder und ihre moralischen Implikationen zu führen, dabei möglichst unter Mitwirkung relevanter Akteure und Interessenvertreter. Auf internationaler Ebene ist es vor allem der Militärbereich, der diesbezüglich die drängendsten Fragen aufwirft (Royakkers/van Est 2015, S.249 ff.) – die Entwicklung teilautonomer Waffensysteme ist bereits weit fortgeschritten und folgt einer militärischen Aufrüstungslogik, die sich nur schwer unter demokratische Kontrolle bringen lässt. Die öffentliche Debatte über die Eingrenzung autonomer Kampfmaschinen,¹²⁸ die sich menschlicher Kontrolle weitgehend entziehen, oder auch die neurotechnologische Beeinflussung von Soldaten (etwa aus therapeutischen Gründen, siehe dazu das SUBNETS-Programm¹²⁹ der DARPA) hat erst vor Kurzem eingesetzt und damit viel zu spät. Dieser Fehler sollte im Pflegebereich, der sich auf nationaler Ebene zu einem der maßgeblichen Anwen-

128 <http://futureoflife.org/open-letter-autonomous-weapons/> (15.12.2015)

129 www.darpa.mil/program/systems-based-neurotechnology-for-emerging-therapies (15.12.2015)



dungsfelder für Technologien der Mensch-Maschine-Entgrenzung entwickeln könnte, nicht wiederholt werden.

Zukünftig ist hierzulande nach allen Prognosen mit einer starken Alterung der Bevölkerung und damit mit einem wachsenden Anteil pflegebedürftiger Menschen an der Gesamtbevölkerung zu rechnen. Der dadurch drohende Pflegenotstand gilt als eine der größten gesellschaftlichen Herausforderungen, die – so wird postuliert – nur mit dem verstärkten Einsatz neuer Technologien zu bewältigen ist. Speziell autonom agierenden Service- bzw. Pflegerobotern sowie der Kombination assistiver Technologien mit nichtinvasiven Neurotechnologien (Steuerung von Exoskeletten, Kommunikationsgeräten etc. mittels Brain-Computer-Interface) wird großes Potenzial zugeschrieben, Pflegekräfte entlasten und Pflegebedürftige im Alltag und bei der Rehabilitation unterstützen zu können – entsprechende Anwendungen befinden sich in Entwicklung und Erprobung. Diesen Chancen stehen jedoch auch ethische Risiken gegenüber, die vor allem im Autonomieverlust der oft nicht mehr artikulationsfähigen und besonders hilfsbedürftigen Pflegebedürftigen zu sehen sind. Kritiker dieser Entwicklung befürchten daher, dass die zukünftige Automatisierung des Pflegebereichs vor allem einer ökonomischen Rationalisierungslogik folgen könnte und nicht mehr den Menschen und seine vielfältigen Bedürfnisse ins Zentrum stellt (Krings et al. 2014, S.96). Kaum zu bestreiten ist, dass der Grat zwischen einem selbstbestimmten Leben durch Technik einerseits und der technikbedingten Entmündigung Pflegebedürftiger andererseits sehr schmal ist. Vor diesem Hintergrund werden zu Recht sozialverträgliche Lösungen gefordert, die sich primär an gesellschaftlichen Bedarfen und menschlichen Bedürfnissen orientieren (Stiftung neue Verantwortung 2013, S.2).

Dass dies nur gemeinsam mit den Betroffenen (Pflegepersonal, ältere Menschen) geht, ergibt sich aus den bisherigen Erörterungen, ebenso, dass die Gestaltung der Pflegeumgebung zentral im Fokus stehen sollte. Wie entsprechende Assistenzsysteme und vor allem auch eine den pflegerischen Herausforderungen angemessene Governance der Technikentwicklung jedoch in der Praxis aussehen könnten, sind Fragen, welche den Rahmen dieses Berichts sprengen und in einem anschließenden TA-Projekt vertieft untersucht werden sollen.





Literatur

In Auftrag gegebene Gutachten

- Behnke, S. (2014): Autonome Robotik, Maschinelles Lernen & Künstliche Intelligenz. Fraunhofer-Institut für Intelligente Analyse- und Informationssysteme (IAIS), Sankt Augustin
- Gransche, B., Warnke, P. (2014): Vision Assessment: Mensch-Maschine-Entgrenzungen. Zwischen künstlicher Intelligenz und Human Enhancement. Fraunhofer-Institut für System- und Innovationsforschung, Karlsruhe
- Hofmann, U. (2014): Das Gehirn im Mittelpunkt. Neuroelektrische Schnittstellen schließen den Regelkreis. Freiburg
- Schaal, S. (2014): Entwicklungen in den Bereichen autonome Robotik, maschinelles Lernen & künstliche Intelligenz. Santa Monica, USA

Weitere Literatur

- Abbott, A. (2012): Mind-controlled robot arms show promise. In: Nature News, www.nature.com/news/mind-controlled-robot-arms-show-promise-1.10652 (4.7.2016)
- acatech (Deutsche Akademie der Technikwissenschaften) (Hg.) (2012): Technikzukünfte. Vorausdenken – Erstellen – Bewerten. acatech IMPULS, Heidelberg, www.acatech.de/fileadmin/user_upload/Baumstruktur_nach_Website/Acatech/root/de/Publikationen/Stellungnahmen/P_2012-11_acatech_Technikzukuenfte_WEB_25102012.pdf (16.6.2016)
- acatech (Hg.) (2016): Innovationspotenziale der Mensch-Maschine-Interaktion. acatech IMPULS, München, www.acatech.de/fileadmin/user_upload/Baumstruktur_nach_Website/Acatech/root/de/Publikationen/Stellungnahmen/acatech_IMPULS_Mensch-Maschine-Interaktion_WEB.pdf (16.6.2016)
- Ahn, M.; Lee, M.; Choi, J.; Jun, S. C. (2014): A review of brain-computer interface games and an opinion survey from researchers, developers and users. In: Sensors 14(8), S. 14601–14633
- Allgeuer, P.; Behnke, S. (2014): Robust sensor fusion for robot attitude estimation. In: Proceedings of 14th IEEE-RAS International Conference on Humanoid Robots (Humanoids), S. 218–224
- Antal, A.; Polania, R.; Schmidt-Samoa, C.; Dechent, P.; Paulus, W. (2011): Transcranial direct current stimulation over the primary motor cortex during fMRI. In: NeuroImage 55(2), S. 590–596
- Bainbridge, W. S. (2006): Appendix 1: Survey of NBIC Applications. In: Bainbridge, W. S.; Roco, M. C. (Hg.): Managing nano-bio-info-cogno innovations. Converging technologies in society. Dordrecht, S. 337–345
- Baldwin, T.; Cole, J.; Fitzgerald, M.; Kitzinger, J.; Laurie, G.; Price, J.; Rose, N.; Rose, S.; Singh, I.; Walsh, V.; Warwick, K. (2013): Novel neurotechnologies: intervening in the brain (Nuffield Council on Bioethics). London



- Beck, S. (2012): Brauchen wir ein Roboterrecht? Ausgewählte juristische Fragen zum Zusammenleben von Menschen und Robotern. In: Japanisch-Deutsches Zentrum (Hg.): Mensch-Roboter-Interaktionen aus interkultureller Perspektive. Japan und Deutschland im Vergleich. Berlin, S. 124–146
- Becker, K.; Börger, S.; Frankenberger, H.; Lüthmann, D.; Norgall, T.; Ohmann, C.; Ranke, A.; Vonthein, R.; Ziegler, A.; Zimolong, A. (2010): Klinische Bewertung, klinische Studien und HTA für Medizinprodukte. Ergebnisbericht zum TMF-Projekt V062-02 »Werkzeuge MP-Entwicklung«. Version 1.2 vom 23.04.2010, www.tmf-ev.de/DesktopModules/Bring2mind/DMX/Download.aspx?Method=attachment&Command=Core_Download&EntryId=8151&PortalId=0 (16.6.2016)
- Beetz, M.; Klank, U.; Kresse, I.; Maldonado, A.; Mösenlechner, L.; Pangercic, D.; Rühr, T.; Tenorth, M. (2011): Robotic roommates making pancakes. In: 11th IEEE-RAS International Conference on Humanoid Robots, S. 529–536
- van den Berg, B. (2011): Robots as tools for techno-regulation. In: *Law, Innovation and Technology* 3(2), S. 317–332
- Berloznik, R.; Casert, R.; Deboelpaep, R.; van Est, R.; Enzing, C.; van Lieshout, M.; Versleijen, A. (2006): Technology assessment on converging technologies (IP/A/STOA/SC/2005-183). Brüssel
- Bernal, J. D. (1970): *The World, the Flesh and the Devil. An enquiry into the future of the three enemies of the rational soul.* London
- Bin, G.; Gao, X.; Yan, Z.; Hong, B.; Gao, S. (2009): An online multi-channel SSVEP-based brain-computer interface using a canonical correlation analysis method. In: *Journal of Neural Engineering* 6(4), S. 46002–46008
- Biran, R.; Martin, D. C.; Tresco, P. A. (2005): Neuronal cell loss accompanies the brain tissue response to chronically implanted silicon microelectrode arrays. In: *Experimental Neurology* 195(1), S. 115–126
- Birbaumer, N. (2000): The thought translation device (TTD) for completely paralyzed patients. In: *IEEE Transactions on Rehabilitation Engineering* 8(2), S. 190–193
- Birbaumer, N. (2006): Breaking the silence: brain-computer interfaces (BCI) for communication and motor control. In: *Psychophysiology* 43(6), S. 517–532
- Birbaumer, N.; Chaudhary, U. (2015): Lernen von Hirnkontrolle – Klinische Anwendung von Brain-Computer Interfaces. In: *Neuroforum* 21(4), S. 130–143
- Bischoff, R.; Kurth, J.; Schreiber, G.; Köppe, R.; Albu-Schäffer, A.; Beyer, A.; Eiberger, O.; Haddadin, S.; Stemmer, A.; Grunwald, G.; Hirzinger, G. (2010): The KUKA-DLR lightweight robot arm – a new reference platform for robotics research and manufacturing. In: Proceedings for the joint conference of ISR 2010 (41st International Symposium on Robotics) und ROBOTIK 2010 (6th German Conference on Robotics). Berlin/Offenbach, S. 741–748
- Bishop, C. (2007): *Pattern recognition and machine learning.* New York
- Böhle, K.; Bopp, K. (2014): What a vision: The artificial companion. A piece of vision assessment including an expert survey. In: *Science, Technology & Innovation Studies* 10(1), S. 155–186
- Bostrom, N. (2014): *Superintelligenz. Szenarien einer kommenden Revolution.* Berlin
- Bregman, A. S. (1994): *Auditory scene analysis: The perceptual organization of sound.* Cambridge/London
- Bronstein, J. M.; Tagliati, M.; Alterman, R. L.; Lozano, A. M.; Volkmann, J.; Stefani, A.; Horak, F. B.; Okun, M. S.; Foote, K. D.; Krack, P.; Pahwa, R. et al. (2011): Deep brain stimulation for Parkinson disease: an expert consensus and review of key issues. In: *Archives of Neurology* 68(2), S. 165–171



- Brunoni, A. R.; Nitsche, M. A.; Bolognini, N.; Bikson, M.; Wagner, T.; Merabet, L.; Edwards, D. J.; Valero-Cabre, A.; Rotenberg, A.; Pascual-Leone, A.; Ferrucci, R. et al. (2012): Clinical research with transcranial direct current stimulation (tDCS): challenges and future directions. In: *Brain Stimulation* 5(3), S. 175–195
- Čapek, K. (1920): R.U.R. (Rossumovi Universální Roboti). Prag
- Carmena, J. M.; Lebedev, M. A.; Crist, R. E.; O'Doherty, J. E.; Santucci, D. M.; Dimitrov, D. F.; Patil, P. G.; Henriquez, C. S.; Nicolelis, M. A. (2003): Learning to control a brain-machine interface for reaching and grasping by primates. In: *Public Library of Science Biology* 1(2), S. 193–208
- Chapin, J. K.; Moxon, K. A.; Markowitz, R. S.; Nicolelis, M. A. L. (1999): Real-time control of a robot arm using simultaneously recorded neurons in the motor cortex. In: *Nature Neuroscience* 2(7), S. 664–670
- Chestek, C. A.; Gilja, V.; Nuyujukian, P.; Foster, J. D.; Fan, J. M.; Kaufman, M. T.; Churchland, M. M.; Rivera-Alvidrez, Z.; Cunningham, J. P.; Ryu, S. I.; Shenoy, K. V. (2011): Longterm stability of neural prosthetic control signals from silicon cortical arrays in rhesus macaque motor cortex. In: *Journal of Neural Engineering* 8(4), S. 45005–45026
- Chrastaller, T.; Decker, M.; Gilsbach, J.; Hirzinger, G.; Lauterbach, K.; Schweighofer, E.; Schweitzer, G.; Sturma, D. (2001): Robotik. Perspektiven für menschliches Handeln in der zukünftigen Gesellschaft. *Wissenschaftsethik und Technikfolgenbeurteilung* 14, Berlin
- Church, G., Regis, E. (2012): *Regenesis. How synthetic biology will reinvent nature and ourselves.* New York
- Churchland, M. M.; Cunningham, J. P.; Kaufman, M. T.; Foster, J. D.; Nuyujukian, P.; Ryu, S. I.; Shenoy, K. V. (2012): Neural population dynamics during reaching. In: *Nature* 487(7405), S. 51–56
- Clark, V. P.; Parasuraman, R. (2014): Neuroenhancement: enhancing brain and mind in health and in disease. In: *NeuroImage* 85, S. 889–894
- Clausen, J. (2010a): Ethical brain stimulation – neuroethics of deep brain stimulation in research and clinical practice. In: *The European Journal of Neuroscience* 32(7), S. 1152–1162
- Clausen, J. (2010b): *Technik im Gehirn: Ethische, theoretische und historische Aspekte moderner Neurotechnologie.* Köln
- Clausen, J. (2012): Ethische Aspekte der Neurostimulation. In: *Das Neurophysiologie-Labor* 34(1–2), S. 68–76
- Clynes, M. E.; Kline, N. S. (1960): Cyborgs and space. In: *Astronautics* 5(9), S. 26–27 und 74–76
- Coenen, C. (2009): Zauberwort Konvergenz. In: *Technikfolgenabschätzung – Theorie und Praxis* 18(2), S. 44–50
- Coenen, C. (2011): Extreme Technikvisionen als Verantwortungsproblem. In: Bartosch, U.; Litfin, G.; Braun, R.; Neuneck, G. (Hg.): *Verantwortung von Wissenschaft und Forschung in einer globalisierten Welt. Forschen – Erkennen – Handeln.* Berlin, S. 231–255
- Coenen, C. (2013): Nachdarwinsche Visionen einer technischen Transformation der Menschheit. In: Ebert, U.; Riha, O.; Zerling, L. (Hg.): *Der Mensch der Zukunft. Hintergründe, Ziele und Probleme des Human Enhancement.* Stuttgart/Leipzig, S. 9–36
- Coenen, C.; Schuijff, M.; Smits, M.; Klaassen, P.; Hennen, L.; Rader, M.; Wolbring, G. (2009): Human enhancement (IP/A/STOA/FWC/2005-28/SC32 & 39). Brüssel



- Coenen, C.; Simakova, E. (2013): STS policy interactions, technology assessment and the governance of technovisionary sciences. In: *Science, Technology & Innovation Studies* 9(2), S. 3–20
- Coffman, B. A.; Clark, V. P.; Parasuraman, R. (2014): Battery powered thought: enhancement of attention, learning, and memory in healthy adults using transcranial direct current stimulation. In: *NeuroImage* 85 Pt 3, S. 895–908
- Colletti, L.; Shannon, R.; Colletti, V. (2012): Auditory brainstem implants for neurofibromatosis type 2. In: *Current opinion in otolaryngology & head and neck surgery* 20(5), S. 353–357
- Collinger, J. L.; Wodlinger, B.; Downey, J. E.; Wang, W.; Tyler-Kabara, E. C.; Weber, D. J. et al. (2013): High-performance neuroprosthetic control by an individual with tetraplegia. In: *The Lancet* 381(9866), S. 557–564
- Collingridge, D. (1982): *The social control of technology*. London
- Davies, S. (2014): Hawking warns on rise of the machines. In: *Financial Times online*, 2.12.2014, www.ft.com/intl/cms/s/2/9943bee8-7a25-11e4-8958-00144feabdc0.html#axzz3ttHVxuIH (16.6.2016)
- Dickow, M. (2015): *Robotik – ein Game-Changer für Militär und Sicherheitspolitik?* Stiftung Wissenschaft und Politik (SWP) – Deutsches Institut für Internationale Politik und Sicherheit, Berlin, www.swp-berlin.org/fileadmin/contents/products/studien/2015_S14_dkw.pdf (16.6.2016)
- Drexler, K. E. (1986): *Engines of creation. The coming era of nanotechnology*. New York
- Dubiel, H. (2006): *Tief im Hirn*. München
- Duret, F.; Brelén, M. E.; Lambert, V.; Gérard, B.; Delbeke, J.; Veraart, C. (2005): Object localization, discrimination, and grasping with the optic nerve visual prosthesis. In: *Restorative neurology and neuroscience* 24(1), S. 31–40
- van Est, R.; Stemerding, D. (Hg.) (2012): *Making perfect life. European governance challenges in 21st century bio-engineering*. Brüssel
- van Est, R. (2014): *Intimate technology. The battle for our body and behaviour*. Den Haag
- Faber, F.; Behnke, S. (2007): Stochastic optimization of bipedal walking using gyro feedback and phase resetting. In: *7th IEEE-RAS International Conference on Humanoid Robots (Humanoids 2007)*, S. 203–209
- Fazli, S.; Mehnert, J.; Steinbrink, J.; Curio, G.; Villringer, A.; Müller, K.-R.; Blankertz, B. (2012): Enhanced performance by a hybrid NIRS–EEG brain computer interface. In: *NeuroImage* 59(1), S. 519–529
- Feil-Seifer, D.; Matarić, M. J. (2005): Defining socially assistive robotics. In: *Proceedings of the IEEE 9th International Conference on Rehabilitation Robotics*, S. 465–468
- Felt, U.; Wynne, B. (2007): *Taking european knowledge society seriously (Report of the expert group on science and governance to the science, economy and society directorate, Directorate-General for Research, European Commission)*. Brüssel, [www.bmbf.de/pub/EuropeanKnowledge\(6\).pdf](http://www.bmbf.de/pub/EuropeanKnowledge(6).pdf) (16.6.2016)
- Ferrari, A.; Coenen, C.; Grunwald, A. (2012): Visions and ethics in current discourse on human enhancement. In: *NanoEthics* 6(3), S. 215–229
- Feuerstein, T. J.; Kammerer, M.; Lücking, C. H.; Moser, A. (2011): Selective GABA release as a mechanistic basis of high-frequency stimulation used for the treatment of neuropsychiatric diseases. In: *Naunyn-Schmiedeberg's archives of pharmacology* 384(1), S. 1–20



- Fraunhofer-Institut für Produktionstechnik und Automatisierung (2015): Roboter als vielseitiger Gentleman. Presseinformation vom 15. Januar 2015, www.care-robot.de/content/dam/careobot/de/documents/Pressemitteilungen/2015_01_13_Care-O-bot_4_final.pdf (16.6.2015)
- Fromme, H. (2015): Ärger für Watson. In: Süddeutsche Zeitung Online, 8.12.2015, www.sueddeutsche.de/wirtschaft/kuenstliche-intelligenz-aerger-fuer-watson-1.2772927 (16.6.2016)
- Gallegos-Ayala, G.; Furdea, A.; Takano, K.; Ruf, C. A.; Flor, H.; Birbaumer, N. (2014): Brain communication in a completely locked-in patient using bed-side near-infrared spectroscopy. In: *Neurology* 82(21), S. 1930–1932
- Gat, E. (1998): On three-layer architectures. In: Kortenkamp, D.; Bonasso, R. P.; Murphy, R. R. (Hg.): *Artificial intelligence and mobile robots*. Cambridge/London, S. 195–210
- Good, I. J. (1965): Speculations concerning the first ultraintelligent machine. In: Alt, F. L.; Rubinoff, M. (Hg.): *Advances in computers*. Bd. 6, New York/London, S. 31–88
- Gopura, R. A. R. C.; Kiguchi, K.; Bandara, D. S. V. (2011): A brief review on up-per extremity robotic exoskeleton systems. In: *Proceedings of 6th International Conference on Industrial and Information Systems*, S. 346–351
- Görz, G.; Schneeberger, J.; Schmid, U. (2014): *Handbuch der künstlichen Intelligenz*. München
- Graimann, B. (2015): Auf dem Weg zur perfekten Prothese. In: *Mensch-Maschine-Visionen: Wie Biologie und Technik verschmelzen*. Spektrum Spezial – Physik, Mathematik, Technik, S. 20–27
- Groppa, S.; Oliviero, A.; Eisen, A.; Quartarone, A.; Cohen, L. G.; Mall, V.; Kaelin-Lang A.; Mima T.; Rossi S.; Thickbroom G. W.; Rossini P. M. et al. (2012): A practical guide to diagnostic transcranial magnetic stimulation: report of an IFCN committee. In: *Clinical Neurophysiology* 123(5), S. 858–882
- Grunwald, A. (2009): Neue Gehirn / Computer-Schnittstellen: Schritte auf dem Weg zur Technisierung des Menschen? In: Hildt, E.; Engels, E.-M. (Hg.): *Der implantierte Mensch. Therapie und Enhancement im Gehirn*. Lebenswissenschaften im Dialog 5, Freiburg/München, S. 183–208
- Grunwald, A. (2011): Einführung in das Schwerpunktthema Hope-, Hype- und Fear-Technologien. In: *TAB-Brief* 39, S. 6–7
- Grunwald, A. (2013): Techno-visionary Sciences: Challenges to policy advice. In: *Science, Technology & Innovation Studies* 9(2), S. 21–38
- Grunwald, A. (2015): Die hermeneutische Erweiterung der Technikfolgenabschätzung. In: *Technikfolgenabschätzung – Theorie und Praxis* 24(2), S. 65–69
- Hairston, W. D.; Whitaker, K. W.; Ries, A. J.; Vettel, J. M.; Bradford, J. C.; Kerick, S. E.; McDowell, K. (2014): Usability of four commercially-oriented EEG systems. In: *Journal of Neural Engineering* 11(4), S. 46018–46032
- Hallett, M. (2000): Transcranial magnetic stimulation and the human brain. In: *Nature* 406(6792), S. 147–150
- Hammond, D. C. (2007): What is neurofeedback? In: *Journal of Neurotherapy* 10(4), S. 25–36
- Haraway, D.J. (1991): A Cyborg manifesto: Science, technology, and socialist-feminism in the late twentieth century. In: Haraway, D.J. (Hg.): *Simians, Cyborgs and women: The reinvention of nature*. New York/London, S. 149–181
- Harrasser, K. (2009): Passung durch Rückkopplung. Konzepte der Selbstregulierung in der Prothetik des Ersten Weltkriegs. In: Fischer, S.; Maehle, E.; Reischuk, R. (Hg.): *Informatik 2009. Im Focus das Leben*. Beiträge der 39. Jahrestagung der Gesellschaft für Informatik e.V., Bonn, S. 788–801



- Hayles, N. K. (1999): How we became posthuman. Virtual bodies in cybernetics, literature, and informatics. Chicago/London
- Hazrati, M. K.; Hofmann, U. G. (2013): Avatar navigation in second life using brain signals. In: 8th IEEE International Symposium on Intelligent Signal Processing, S. 1–7
- Healy, P.; Rayner, S. (Hg.) (2009): Unnatural selection: The challenges of engineering tomorrow's people. London/Sterling
- Heck, C. N.; King-Stephens, D.; Massey, A. D.; Nair, D. R.; Jobst, B. C.; Barkley, G. L.; Salanova, V.; Cole, A. J.; Smith, M. C.; Gwinn, R. P.; Skidmore, C. et al. (2014): Two-year seizure reduction in adults with medically intractable partial onset epilepsy treated with responsive neurostimulation: Final results of the RNS system pivotal trial. In: *Epilepsia* 55(3), S. 432–441
- Heffley, H. (2010): Pediatric cochlear implants: Medical miracle or cultural genocide? In: *The Journal of Undergraduate Nursing Writing* 4(1), S. 17–23
- Heil, R. (2010a): Trans- und Posthumanismus. Eine Begriffsbestimmung. In: Hilt, A.; Frewer, A.; Jordan, I. (Hg.): *Endlichkeit, Medizin und Unsterblichkeit. Geschichte – Theorie – Ethik*, Stuttgart, S. 127–150
- Heil, R. (2010b): Human Enhancement – Eine Motivsuche bei J.D. Bernal, J.B.S. Haldane und J. Huxley. In: Coenen, C.; Gammel, S.; Heil, R.; Woyke, A. (Hg.): *Die Debatte über »Human Enhancement«*. Historische, philosophische und ethische Aspekte der technologischen Verbesserung des Menschen. *Science Studies*, Bielefeld, S. 41–62
- Heil, R.; Coenen, C. (2014): Transhumanistische Visionen und Computertechnik. In: Böhle, K.; Berendes, J.; Gutmann, M.; Robertson-von Trotha, C.; Scherz, C. (Hg.): *Computertechnik und Sterbekultur. Hermeneutik und Anthropologie* 5, Berlin/Münster, S. 145–159
- Hennen, L.; Grünwald, R.; Revermann, C.; Sauter, A. (2008): Einsichten und Eingriffe in das Gehirn. Die Herausforderung der Gesellschaft durch die Neurowissenschaften. *Studien des Büros für Technikfolgen-Abschätzung* 24, Berlin
- Herff, C.; Heger, D.; de Pestors, A.; Telaar, D.; Brunner, P.; Schalk, G.; Schultz, T. (2015): Brain-to-text: decoding spoken phrases from phone representations in the brain. In: *Frontiers in Neuroscience* 9, Art. 217
- Hertz, J. A.; Krogh, A. S.; Palmer, R. G. (1991): *Introduction to the theory of neural computation*. Santa Fe Institute Studies in the Sciences of Complexity 1, Boulder
- Herwig, U.; Fallgatter, A. J.; Höppner, J.; Eschweiler, G. W.; Kron, M.; Hajak, G.; Padberg F.; Naderi-Heiden A.; Abler B.; Eichhammer P.; Grossheinrich N. et al. (2007): Antidepressant effects of augmentative transcranial magnetic stimulation: randomised multicentre trial. In: *The British Journal of Psychiatry: the Journal of Mental Science* 191, S. 441–448
- Hild, M.; Siedel, T.; Benckendorff, C.; Thiele, C.; Spranger, M. (2012): Myon, a new humanoid. In: Steels, L.; Hild, M. (Hg.): *Language grounding in robots*. New York u. a. O., S. 25–44
- Hilgendorf, E. (2012): Können Roboter schuldhaft handeln? Zur Übertragbarkeit unseres normativen Grundvokabulars auf Maschinen. In: Beck, S. (Hg.): *Jenseits von Mensch und Maschine: Ethische und rechtliche Fragen zum Umgang mit Robotern, künstlicher Intelligenz und Cyborgs*. Baden-Baden, S. 119–132
- Hinton, G. E.; Salakhutdinov, R. R. (2006): Reducing the dimensionality of data with neural networks. In: *Science* 313(5786), S. 504–507
- Hirsch-Kreinsen, H. (2014): Wandel von Produktionsarbeit–»Industrie 4.0«. *Soziologisches Arbeitspapier* 38, www.wiso.tu-dortmund.de/wiso/ts/de/forschung/veroeff/soz_arbeitspapiere/AP-SOZ-38.pdf (17.6.2016)



- Hochberg, L. R.; Bacher, D.; Jarosiewicz, B.; Masse, N. Y.; Simeral, J. D.; Vogel, J.; Haddadin, S.; Liu, J.; Cash, S. S.; van der Smagt, P.; Donoghue, J. P. (2012): Reach and grasp by people with tetraplegia using a neurally controlled robotic arm. In: *Nature* 485(7398), S. 372–375
- Hochberg, L. R.; Serruya, M. D.; Friehs, G. M.; Mukand, J. A.; Saleh, M.; Caplan, A. H.; Branner, A.; Chen, D.; Penn, R. D.; Donoghue, J. P. (2006): Neuronal ensemble control of prosthetic devices by a human with tetraplegia. In: *Nature* 442(7099), S. 164–171
- Humayun, M. S.; Dorn, J.D.; da Cruz, L.; Dagnelie, G.; Sahel, J.-A.; Stanga, P. E.; Cideciyan A. V.; Duncan, J. L.; Elliott, D.; Filley, E.; Ho, A. C. et al. (2012): Interim results from the international trial of second sight's visual prosthesis. In: *Ophthalmology* 119(4), S. 779–788
- Huster, R. J.; Mokom, Z. N.; Enriquez-Geppert, S.; Herrmann, C. S. (2014): Brain-computer interfaces for EEG neurofeedback: peculiarities and solutions. In: *International Journal of Psychophysiology: official Journal of the International Organization of Psychophysiology* 91(1), S. 36–45
- Huxley, A. (1932): *Brave New World*. London
- ICRC (International Committee of the Red Cross) (2014): *Autonomous weapon systems. Technical, military, legal and humanitarian aspects*. 26.–28.3.2014 Genf, www.icrc.org/en/download/file/1707/4221-002-autonomous-weapons-systems-full-report.pdf (17.6.2016)
- IFR (International Federation of Robotics) (2015a): *World Robotics 2015: Industrial Robots*. Frankfurt am Main
- IFR (2015b): *World Robotics 2015: Service Robots*. Frankfurt am Main
- Jansen, M. (2015): *Digitale Herrschaft: Über das Zeitalter der globalen Kontrolle und wie Transhumanismus und Synthetische Biologie das Leben neu definieren*. Stuttgart
- Jasanoff, S.; Kim, S-H. (2009): Containing the atom: Sociotechnical imaginaries and nuclear power in the United States and South Korea. In: *Minerva* 47, S. 119–146
- Kaiser, M. (2015): Reactions to the future: the chronopolitics of prevention and preemption. In: *NanoEthics* 9(2), S. 165–177
- Kaminski, E.; Hoff, M.; Sehm, B.; Taubert, M.; Conde, V.; Steele, C. J.; Villringer, A.; Ragert, P. (2013): Effect of transcranial direct current stimulation (tDCS) during complex whole body motor skill learning. In: *Neuroscience Letters* 552, S. 76–80
- Kaplan, A.; Shishkin, S. L.; Ganin, I. P.; Basyul, I. A.; Zhigalov, A. Y. (2013): Adapting the P300-based brain-computer interface for gaming: a review. In: *IEEE transactions on computational intelligence and AI in games* 5(2), S. 141–149
- Kelly, S. P.; Lalor, E. C.; Reilly, R. B.; Foxe, J. J. (2005): Visual spatial attention tracking using high-density SSVEP data for independent brain-computer communication. In: *IEEE transactions on neural systems and rehabilitation engineering* 13(2), S. 172–178
- Kläß, J.; Stückler, J.; Behnke, S. (2012): Efficient mobile robot navigation using 3D surfel grid maps. In: *Proceedings for the conference of ROBOTIK 2012, 7th German Conference on Robotics*, S. 1–4
- Kotler, S. (2002): *Vision Quest*. In: *Wired Magazine*, 9.1.2002, www.wired.com/2002/09/vision/ (6.7.2016)
- Krings, B.-J.; Böhle, K.; Decker, M.; Nierling, L.; Schneider, C. (2014): *Serviceroboter in Pflegearrangements*. In: Decker, M.; Fleischer, T.; Schippl, J.; Weinberger, N. (Hg.): *Zukünftige Themen der Innovations- und Technikanalyse: Lessons learned und ausgewählte Ergebnisse*. Karlsruhe, S. 63–121



- Kübler, A.; Mattia, D.; Rupp, R.; Tangermann, M. (2013): Facing the challenge: bringing brain-computer interfaces to end-users. In: *Artificial Intelligence in Medicine* 59(2), S. 55–60
- Kübler, A.; Neuper, C. (2008): Gehirn-Computer-Schnittstellen (Brain-Computer-Interfaces): Anwendungen und Perspektiven. In: *Neuroforum* 14(2), S. 204–210
- Kuhn, J.; Gründler, T. O. J.; Lenartz, D.; Sturm, V.; Klosterkötter, J.; Huff, W. (2010): Tiefe Hirnstimulation bei psychiatrischen Erkrankungen. In: *Deutsches Ärzteblatt International* 107(7), S. 105–113
- Kuiken, T. A.; Li, G.; Lock, B. A.; Lipschutz, R. D.; Miller, L. A.; Stubblefield, K. A.; Englehart, K. B. (2009): Targeted muscle reinnervation for real-time myoelectric control of multifunction artificial arms. In: *Jama* 301(6), S. 619–628
- Kurzweil, R. (2005): *The singularity is near: When humans transcend biology*. New York
- Lampe, T.; Fiederer, L. D.J.; Völker, M.; Knorr, A.; Riedmiller, M.; Ball, T. (2014): A brain-computer interface for high-level remote control of an autonomous, reinforcement-learning-based robotic system for reaching and grasping. In: *Proceedings of the 19th international conference on Intelligent User Interfaces*. New York, S. 83–88
- Lanier, J. (2014): *The Myth of AI. A conversation with Jaron Lanier*. www.edge.org/conversation/jaron_lanier-the-myth-of-ai (16.6.2016)
- Laxton, A. W.; Lipsman, N.; Lozano, A. M. (2013): Deep brain stimulation for cognitive disorders. In: *Handbook of Clinical Neurology* 116, S. 307–311
- Licklider, J. C. R. (1960): Man-computer symbiosis. In: *IRE Transactions on Human Factors in Electronics* (Bd. HFE-1), S. 4–11
- Lösch, A. (2013): »Vision Assessment« zu Human-Enhancement-Technologien. Konzeptionelle Überlegungen zu einer Analytik von Visionen im Kontext gesellschaftlicher Kommunikationsprozesse. In: *Technikfolgenabschätzung – Theorie und Praxis* 22(1), S. 9–16
- Lotte, F.; Bougrain, L.; Clerc, M. (2015): Electroencephalography (EEG)-based brain-computer interfaces. In: *Wiley Encyclopedia of Electrical and Electronics Engineering*, S. 1–44
- Lozano, A. M.; Lipsman, N. (2013): Probing and regulating dysfunctional circuits using deep brain stimulation. In: *Neuron* 77(3), S. 406–424
- Mackenzie, R. (2011): Who should hold the remote for the new me? Cognitive, affective and behavioural side-effects of DBS and authentic choices over future personalities. In: *AJOB Neuroscience* 2(1), S. 18–20
- Mannino, A.; Althaus, D.; Erhardt, J.; Gloor, L.; Hutter, A.; Metzinger, T. (2015): Künstliche Intelligenz: Chancen und Risiken. *Diskussionspapiere der Stiftung für Effektiven Altruismus* 2, S. 1–17
- Maslen, H.; Douglas, T.; Cohen Kadosch, R.; Levy, N.; Savulsecu, J. (2014): The regulation of cognitive enhancement devices: extending the medical model. In: *Journal of Law and the Biosciences* 1(1), S. 68–93
- Mavroidis, C.; Ferreira, A. (2013): Nanorobotics: Past, present, and future. In: Mavroidis, C.; Ferreira, A. (Hg.): *Nanorobotics. Current approaches and techniques*. New York, S. 3–27
- McCray, W. P. (2012): *The visioneers: How a group of elite scientists pursued space colonies, nanotechnologies, and a limitless future*. Princeton/Oxford
- Medtronic (2015): Medtronic gibt europaweite Zulassung der ersten und einzigen Ganzkörper-MRT-fähigen Systeme für die tiefe Hirnstimulation bekannt. Pressemitteilung vom 21.04.2015, www.medtronic.de/wcm/groups/mdtcom_sg/@mdt/@eu/@de/documents/documents/ce_mark_mrisafe_dbs.pdf (07.03.2016)



- Meinzer, M.; Jähnigen, S.; Copland, D. A.; Darkow, R.; Grittner, U.; Avirame, K.; Rodriguez, A. D.; Lindenberg, R.; Flöel, A. (2014): Transcranial direct current stimulation over multiple days improves learning and maintenance of a novel vocabulary. In: *Cortex* 50, S. 137–147
- Metta, G.; Sandini, G.; Vernon, D.; Natale, L.; Nori, F.; (2008): The iCub humanoid robot: an open systems platform for research in embodied cognition. In: Madhavan, R. (Hg.): *Proceedings of the 8th Workshop on Performance Metrics for Intelligent Systems*. New York, S. 50–56
- Meyer-Drawe, K. (1996): *Menschen im Spiegel ihrer Maschinen. Übergänge. Texte zu Handlung, Sprache und Lebenswelt* 29, München
- del R. Millán, J.; Rupp, R.; Müller-Putz, G. R.; Murray-Smith, R.; Giugliemma, C.; Tangermann, M.; Vidaurre, C.; Cincotti, F.; Kübler, A.; Leeb, R.; Neuper, C.; Müller, K.-R. (2010): Combining brain-computer interfaces and assistive technologies: state-of-the-art and challenges. In: *Frontiers in Neuroscience* 4(161), S. 1–16
- del R. Millán, J.; Carmena, J. M. (2010): Invasive or noninvasive: understanding brain-machine interface technology. In: *IEEE Engineering in Medicine and Biology Magazine* 29(1), S. 16–22
- Miranda, R. A.; Casebeer, W. D.; Hein, A. M.; Judy, J. W.; Krotkov, E. P.; Laabs, T. L.; Manzo, J. E.; Pankratz, K. G.; Pratt, G. A.; Sanchez, J. C.; Weber, D. J.; Wheeler, T. L. (2015): DARPA-funded efforts in the development of novel brain-computer interface technologies. In: *Journal of Neuroscience Methods* 244, S. 52–67
- Monyer, H.; Rösler, F.; Roth, G.; Scheich, H.; Singer, W.; Elger, C. E.; Friederici, A. D.; Koch, C.; Luhmann, H.; von der Malsburg, C.; Menzel, R. (2004): Das Manifest. Elf führende Neurowissenschaftler über Gegenwart und Zukunft der Hirnforschung. In: *Gehirn und Geist* 6, S. 30–37
- Moravec, H. (2001): *Mind children. Der Wettlauf zwischen menschlicher und künstlicher Intelligenz*. Hamburg
- Müller, M. F. (2014): Roboter und Recht. Eine Einführung. In: *Aktuelle Juristische Praxis* (5), S. 595–608
- Müller, O. (2010): *Zwischen Mensch und Maschine: vom Glück und Unglück des Homo faber*. Edition Unselde 29, Frankfurt am Main
- Müller, S.; Zaracko, A. (2010): Haben gehörlose Kleinkinder ein Recht auf ein Cochleaimplantat? In: *Nervenheilkunde* 29(4), S. 244–248
- Neuhäuser, C. (2012): Künstliche Intelligenz und ihr moralischer Standpunkt. In: Beck, S. (Hg.): *Jenseits von Mensch und Maschine: Ethische und rechtliche Fragen zum Umgang mit Robotern, künstlicher Intelligenz und Cyborgs*. Baden-Baden, S. 23–42
- Nguyen, J. K.; Park, D. J.; Skousen, J. L.; Hess-Dunning, A. E.; Tyler, D. J.; Rowan, S. J.; Weder, C.; Capadona, J. R. (2014): Mechanically-compliant intra-cortical implants reduce the neuroinflammatory response. In: *Journal of Neural Engineering* 11(5), S. 056014
- Nicolelis, M. A. L.; Chapin, J. K. (2008): Controlling robots with the mind. In: *Scientific American* 18, S. 72–79
- Niedermeyer, E.; Lopes da Silva, F. F. H. (Hg.) (2005): *Electroencephalography: basic principles, clinical applications, and related fields*. Philadelphia
- Nieuwenhuisen, M.; Behnke, S. (2014): Layered mission and path planning for MAV navigation with partial environment knowledge. In: *Proceedings of the International Conference on Intelligent Autonomous Systems (IAS)*, S. 1–13
- Nieuwenhuisen, M.; Behnke, S. (2013): Human-like interaction skills for the mobile communication robot Robotinho. In: *International Journal of Social Robotics* 5(4), S. 549–561



- Nitsche, M. A.; Paulus, W. (2000): Excitability changes induced in the human motor cortex by weak transcranial direct current stimulation. In: *The Journal of Physiology* 527(3), S. 633–639
- Nordmann, A. (2007): If and Then: A Critique of Speculative NanoEthics. In: *NanoEthics* 1(1), S. 31–46
- Nordmann, A. (2013): Visioneering assessment: on the construction of tunnel visions for technovisionary research and policy. In: *Science, Technology & Innovation Studies* 9(2), S. 89–94
- Normann, R. A.; Maynard, E. M.; Rousche, P. J.; Warren, D. J. (1999): A neural interface for a cortical vision prosthesis. In: *Vision Research* 39(15), S. 2577–2587
- Nuyujukian, P.; Kao, J. C.; Fan, J. M.; Stavisky, S. D.; Ryu, S. I.; Shenoy, K. V. (2014): Performance sustaining intracortical neural prostheses. In: *Journal of Neural Engineering* 11(6), S. 066003
- Oberhofer, W.; Zimmerer, T. (1996): Wie Künstliche Neuronale Netze lernen: Ein Blick in die Black Box der Backpropagation Netzwerke. In: *Regensburger Diskussionsbeiträge* (287), S. 1–46
- O'Reardon, J. P.; Solvason, H. B.; Janicak, P. G.; Sampson, S.; Isenberg, K. E.; Nahas, Z.; McDonald, W. M.; Avery, D.; Fitzgerald, P. B.; Loo, C.; Demitrack, M. A.; et al. (2007): Efficacy and safety of transcranial magnetic stimulation in the acute treatment of major depression: a multisite randomized controlled trial. In: *Biological Psychiatry. A Journal of Psychiatric Neuroscience and Therapeutics* 62(11), S. 1208–1216
- Ortiz-Catalan, M. J.; Brånemark, R.; Håkansson, B.; Delbeke, J. (2012): On the viability of implantable electrodes for the natural control of artificial limbs: review and discussion. In: *Biomedical Engineering Online* 11(33), S. 1–24
- Park, E. (2014): Ethical issues in cyborg technology: Diversity and inclusion. In: *NanoEthics* 8(3), S. 303–306
- Park, I.-W.; Kim, J.-Y.; Lee, J.; Oh, J.-H. (2007): Mechanical design of the humanoid robot platform, HUBO. In: *Advanced Robotics* 21(11), S. 1305–1322
- Paschen, H.; Coenen, C.; Fleischer, T.; Grünwald, R.; Oertel, D.; Revermann, C. (2004): *Nanotechnologie. Forschung, Entwicklung, Anwendung*. Berlin/Heidelberg
- Pearlman, E. (2015): I, cyborg. In: *PAJ: A Journal of Performance and Art* 37(2), S. 84–90
- Petit, D.; Gergondet, P.; Cherubini, A.; Kheddar, A. (2015): An integrated framework for humanoid embodiment with a BCI. In: *IEEE International Conference on Robotics and Automation (ICRA)*, S. 2882–2887
- Pfeifer, R.; Bongard, J. (2007): *How the body shapes the way we think. A new view of intelligence*. Cambridge/London
- Pfurtscheller, G.; Neuper, C.; Flotzinger, D.; Pregenzer, M. (1997): EEG-based discrimination between imagination of right and left hand movement. In: *Electroencephalography and Clinical Neurophysiology* 103(6), S. 642–651
- Pfurtscheller, G.; Neuper, C. (2001): Motor imagery and direct brain-computer communication. In: *Proceedings of the IEEE* 82(7), S. 1123–1134
- Pham, U. H. G.; Solbakk, A.-K.; Skogseid, I.-M.; Toft, M.; Pripp, A. H.; Eidahl Konglund, A. E.; Andersson, S.; Haraldsen, I. R.; Aarsland, D.; Dietrichs, E.; Malt, U. F. (2015): Personality Changes after Deep Brain Stimulation in Parkinson's Disease. In: *Parkinsons Disease* 29, S. 1–7
- Quante, M. (2014): *Personale Identität und Tiefenhirnstimulation*. In: *Preprints and Working Papers of the Centre for Advanced Study in Bioethics* (44), S. 1–20
- Rabiner, L.; Juang, B.-H. (1993): *Fundamentals of speech recognition*. Englewood Cliffs



- Raspopovic, S.; Capogrosso, M.; Petrini, F. M.; Bonizzato, M.; Rigosa, J.; di Pino, G.; Carpaneto, J.; Controzzi, M.; Boretius, T.; Fernandez, E.; Granata, G. et al. (2014): Restoring natural sensory feedback in real-time bidirectional hand prostheses. In: *Science Translational Medicine* 6(222), S. 222ra19
- RoboLaw (Regulating Emerging Robotic Technologies in Europe: Robotics facing Law and Ethics) (2014): Guidelines on Regulating Robotics (D 6.2). www.robotlaw.eu (6.7.2016)
- Roco, M. C.; Bainbridge, W. S. (Hg.) (2003). *Converging technologies for improving human performance. Nanotechnology, biotechnology, information technology and cognitive science.* Dordrecht
- Royakkers, L.; van Est, R. (2015): *Just ordinary robots: automation from love to war.* Boca Raton/London/New York
- Rucci, M.; Bullock, D.; Santini, F. (2007): Integrating robotics and neuroscience. Brains for robots, bodies for brains. In: *Advanced Robotics* 21(10), S. 1115–1129
- Rumelhart, D. E.; McClelland, J. L. (1986): *Parallel distributed processing: Explorations in the microstructure of cognition. Volume 1: Foundations,* Cambridge
- Russell, S. J.; Norvig, P. (2014): *Artificial intelligence. A modern approach.* Harlow
- Saage, R. (2011): *Philosophische Anthropologie und der technisch aufgerüstete Mensch. Annäherungen an Strukturprobleme des biologischen Zeitalters.* Bochum
- Sale, P.; Franceschini, M.; Waldner, A.; Hesse, S. (2012): Use of the robot assisted gait therapy in rehabilitation of patients with stroke and spinal cord injury. In: *European Journal of Physical and Rehabilitation Medicine* 48(1), S. 111–121
- Sanchez, J. C.; Alba, N.; Nishida, T.; Batich, C.; Carney, P. R. (2006): Structural modifications in chronic microwire electrodes for cortical neuroprosthetics: a case study. In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 14(2), S. 217–221
- Sauter, A.; Gerlinger, K. (2012): *Der pharmakologisch verbesserte Mensch. Leistungssteigernde Mittel als gesellschaftliche Herausforderung. Studien des Büros für Technikfolgen-Abschätzung* 34, Berlin
- Schaal, S. (2007): The new robotics – towards human-centered machines. In: *HFSP Journal* 1(2), S. 115–126
- Schalk, G.; McFarland, D. J.; Hinterberger, T.; Birbaumer, N.; Wolpaw, J. R. (2004): BCI2000: a general-purpose brain-computer interface (BCI) system. In: *IEEE Transactions on Biomedical Engineering* 51(6), S. 1034–1043
- Schermer, M. (2011): Ethical issues in deep brain stimulation. In: *Frontiers in Integrative Neuroscience* 5(17), S. 1–5
- Schirmacher, F. (Hg.) (2001): *Die Darwin AG. Wie Nanotechnologie, Biotechnologie und Computer den neuen Menschen träumen.* Köln
- Schneider, S. (2009): Future minds: Transhumanism, cognitive enhancement, and the nature of persons. In: Ravitsky, V.; Fiester, A.; Caplan, A. L. (Hg.): *The Penn Center Guide to Bioethics.* New York, S. 95–100
- Schneider, E.; Strauß, G. (2013): Training der Selbstkontrolle der langsamen kortikalen Potenziale. In: Haus, K.-M.; Held, C.; Kowalski, A.; Krombholz, A.; Nowak, M.; Schneider, E.; Strauß, E.; Wiedemann, M. (Hg.): *Praxisbuch Biofeedback und Neurofeedback.* Berlin/Heidelberg, S. 61–88
- Schneider, C.; Lösch, A. (2015): What about your futures, Technology Assessment? An essay on how to take the visions of TA seriously, motivated by the PACITA conference. In: *Technikfolgenabschätzung – Theorie und Praxis* 24(2), S. 70–74
- von Schomberg, René (2014): From ›Responsible Development of Technologies‹ to Responsible Innovation. <http://renevonschomberg.wordpress.com/from-responsible-development-of-technologies-to-responsible-innovation/> (6.7.2016)



- Schreuder, M.; Rost, T.; Tangermann, M. (2011): Listen, you are writing! Speeding up online spelling with a dynamic auditory BCI. In: *frontiers in Neuroscience* 5(112), S. 1121–12
- Schuepbach, W. M. M.; Rau, J.; Knudsen, K.; Volkmann, J.; Krack, P.; Timmermann, L.; Hälbig, T. D.; Hesekamp, H.; Navarro, S. M.; Meier, N.; Falk, D. et al. (2013): Neurostimulation for Parkinson's disease with early motor complications. In: *New England Journal of Medicine* 368(7), S. 610–622
- Schulz, T. (2015): »Ich bin einfach Optimist« (Interview mit Google-Gründer Larry Page). In: *DER SPIEGEL* 43, S. 104–110
- Schummer, J. (2009): *Nanotechnologie: Spiele mit Grenzen*. Edition Unseld 23, Frankfurt am Main
- Schwartz, A. B. (2004): Cortical neural prosthetics. In: *Annual Review of Neuroscience* 27, S. 487–507
- Schwarz, M.; Stückler, J.; Behnke, S. (2014): Mobile teleoperation interfaces with adjustable autonomy for personal service robots. In: *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction*, S. 288–289
- Selin, C. (2008): The sociology of the future: Tracing stories of technology and time. In: *Sociology Compass* 2(6), S. 1878–1895.
- Serruya, M. D. (2014): Bottlenecks to clinical translation of direct brain-computer interfaces. In: *Frontiers in Systems Neuroscience* 8(Art. 226) S. 1–6
- Service, R. F. (2013): The cyborg era begins. In: *Science* 340(6137), S. 1162–1165
- Seung, S. (2013): *Das Konnektom – Erklärt der Schaltplan des Gehirns unser Ich?* Berlin/Heidelberg
- Shelley, M. (1818): *Frankenstein or The Modern Prometheus*. London
- Shih, J. J.; Krusienski, D. J.; Wolpaw, J. R. (2012): Brain-computer interfaces in medicine. In: *Mayo Clinic Proceedings* 87(3), S. 268–279
- Siegbahn, M.; Lundin, K.; Olsson, G.-B.; Stillesjö, F.; Kinnefors, A.; Rask-Andersen, H.; Nyberg, G. (2014): Auditory brainstem implants (ABIs) – 20 years of clinical experience in Uppsala, Sweden. In: *Acta Otolaryngologica* 134(10), S. 1052–1061
- Silver, D.; Huang, A.; Maddison, C. J.; Guez, A.; Sifre, L.; van den Driessche, G.; Schrittwieser, J.; Antonoglou, I.; Panneershelvam, V.; Lanctot, M.; Dieleman, S. et al. (2016): Mastering the game of Go with deep neural networks and tree search. In: *Nature* 529(7587), S. 484–489
- Silver, J.; Miller, J. H. (2004): Regeneration beyond the glial scar. In: *Nature Reviews Neuroscience* 5(2), S. 146–156
- Simakova, E.; Coenen, C. (2013): Visions, hype, and expectations: A place for responsibility. In: Owen, R.; Heintz, M.; Bessant, J. (Hg.): *Responsible innovation: Managing the responsible emergence of science and innovation in society*. Chichester, S. 241–267
- Slusser, G. (2009): Dimorphs and doubles: J. D. Bernal's »Two Cultures« and the transhuman promise. In: Westfahl, G.; Slusser, G. (Hg.): *Science Fiction and the two cultures: Essays on bridging the gap between the sciences and the humanities*. Jefferson, S. 96–130
- Spatz, J. P.; Schaal, S. (2014): Intelligent systems: Bits and bots. In: *Nature Supplements* 509(7502), S. 1–7
- Stieglitz, T. (2015): Neuroimplantate. In: *Mensch Maschine Visionen. Wie Biologie und Technik verschmelzen: Spektrum der Wissenschaft (Spektrum spezial, 13.5.2015)*, S. 6–13
- Stieglitz, T.; Hofmann, U. G.; Rosahl, S. (Hg.) (2014): *Neurotechnik*. In: Morgenstern, U.; Kraft, M. (Hg.): *Biomedizinische Technik – Faszination, Einführung, Überblick*. Band 1, Berlin/Boston, S. 441–466



- Stiftung neue Verantwortung (2013): Mit Robotern gegen den Pflegenotstand. Policy Brief, S. 1–11
- Stingl, K.; Bartz-Schmidt, K. U.; Besch, D.; Braun, A.; Bruckmann, A.; Gekeler, F.; Greppmaier, U.; Hipp, S.; Hörtdörfer, G.; Kernstock, C.; Koitschev, A. et al. (2013): Artificial vision with wirelessly powered subretinal electronic implant alpha-IMS. In: *Proceedings of the Royal Society B. Biological Sciences* 280(1757), S. 20130077
- Stückler, J.; Behnke, S. (2014): Adaptive tool-use strategies for anthropomorphic service robots. In: *14th IEEE-RAS International Conference on Humanoid Robots (Humanoids 2014)*, S. 755–760
- Stückler, J.; Holz, D.; Behnke, S. (2012): Demonstrating everyday manipulation skills in RoboCup@Home. In: *IEEE Robotics and Automation Magazine* 19(2), S. 34–42
- Sturma, D. (2003): Autonomie. Über Personen, künstliche Intelligenz und Robotik. In: Christaller, T.; Wehner, J. (Hg.): *Autonome Maschinen*. Wiesbaden, S. 38–55
- Suner, S.; Fellows, M. R.; Vargas-Irwin, C.; Nakata, G. K.; Donoghue, J. P. (2005): Reliability of signals from a chronically implanted, silicon-based electrode array in non-human primate primary motor cortex. In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 13(4), S. 524–541
- Sung, J.-Y.; Guo, L.; Grinter, R. E.; Christensen, H. I. (2007): »My Roomba Is Rambo«: Intimate home appliances. In: Krumm, J.; Abowd, G. D.; Seneviratne, A.; Strang, T. (Hg.): *UbiComp 2007: ubiquitous computing. 9th International Conference, UbiComp 2007, Innsbruck, Austria, September 16–19, 2007, Proceedings. Lecture Notes in Computer Science* 4717, Berlin/Heidelberg/New York, S. 145–162
- Sutton, R. S.; Barto, A. G. (1998): *Introduction to reinforcement learning*. Cambridge
- Suzuki, K. (Hg.) (2013): *Artificial neural networks – architectures and applications*. Rijeka
- Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. et al. (2014): Going deeper with convolutions. In: *arXiv:1409.4842*, S. 1–12
- TAB (Büro für Technikfolgen-Abschätzung beim Deutschen Bundestag) (2008): *Konvergierende Technologien und Wissenschaften. Der Stand der Debatte und politischen Aktivitäten zu »converging technologies«* (Autor: Coenen, C.). TAB-Hintergrundpapier Nr. 16, Berlin
- TAB (2011): *Stand und Perspektiven der militärischen Nutzung unbemannter Systeme* (Autoren: Petermann, T., Grünwald, R.). TAB-Arbeitsbericht Nr. 144, Berlin
- Tabbert, T. T. (2004): *Menschmaschinen-götter. Künstliche Menschen in Literatur und Technik. Fallstudien einer Artifzianthropologie*. Hamburg
- Takahashi, T. (2010): *Kybernetik in Japan – Roboter und Cyborg. Beitrag auf dem Symposium »Mensch-Roboter-Interaktionen aus interkultureller Perspektive: Japan und Deutschland im Vergleich«*, 7.–8.12.2010, Berlin
- Tanriverdi, H. (2015): So will Elon Musk künstliche Intelligenz bändigen. In: *Süddeutsche Zeitung Online*, 12.12.2015
- Thrun, S. B. (1992): The role of exploration in learning control. In: White, D. A.; Sofge, D. A. (Hg.): *Handbook of intelligent control. Neural, fuzzy, and adaptive approaches*. New York, S. 527–559
- Tirosh-Samuels, H. (2012): Science and the betterment of humanity. Three british prophets of transhumanism. In: Tirosh-Samuels, H.; Mossman, K. L. (Hg.): *Building better humans? Refocusing the debate on transhumanism. Band 3 von Beyond humanism: Trans- and posthumanism*, Frankfurt am Main u. a. O., S. 55–82
- Tottori, S.; Zhang, L.; Nelson, B. J. (2014): Wireless Actuation of Micro/Nanorobots for Medical Applications. In: Ge, Y.; Li, S.; Wang, S.; Moore, R. (Hg.): *Nanomedicine. Principles and perspectives*. New York, S. 171–189



- Toussaint, M.; Ritter, H.; Jost, J.; Igel, C. (2010): Antrag auf Einrichtung eines neuen DFG-Schwerpunktprogramms Autonomes Lernen (Kurzversion). <http://ipvs.informatik.uni-stuttgart.de/mlr/spp-wordpress/wp-content/uploads/antrag-public.pdf> (12.7.2016)
- Tronnier, V. M.; Fogel, W.; Krause, M.; Bonsanto, M. M.; Tronnier, J.; Heck, A.; Münkkel, K.; Kunze, S. (2002): High frequency stimulation of the basal ganglia for the treatment of movement disorders: Current status and clinical results. In: *Minimally invasive neurosurgery* 45(2), S. 91–96
- VDMA (Verband Deutscher Maschinen- und Anlagenbau) (2014): Positionspapier: Sicherheit bei der Mensch-Roboter-Kollaboration. www.dira.dk/media/66065/sikkerhed_ved_menneske-robot_kollaboration.pdf (12.7.2016)
- Vinyals, O.; Toshev, A.; Bengio, S.; Erhan, D. (2014): Show and tell: A neural image caption generator. In: arXiv:1411.4555, S. 1–9
- Wada, K.; Shibata, T. (2007): Living with seal robots – its sociopsychological and physiological influences on the elderly at a care house. In: *IEEE Transactions on Robotics* 23(5), S. 972–980
- Wagner, T. (2015): *Robokratie: Google, das Silicon Valley und der Mensch als Auslaufmodell*. Köln
- Wainer, J.; Dautenhahn, K.; Robins, B.; Amirabdollahian, F. (2010): Collaborating with Kaspar: Using an autonomous humanoid robot to foster cooperative dyadic play among children with autism. In: 10th IEEE-RAS International Conference on Humanoid Robots (Humanoids 2010), S. 631–638
- Waldrop, M. M. (2013): Neuroelectronics: Smart connections. In: *Nature* 503(7474), S. 22–24
- Walters, M. L.; Syrdal, D. S.; Dautenhahn, K.; te Boekhorst, R.; Koay, K. L. (2008): Avoiding the uncanny valley: robot appearance, personality and consistency of behavior in an attention-seeking home scenario for a robot companion. In: *Autonomous Robots* 24(2), S. 159–178
- Wessberg, J.; Stambaugh, C. R.; Kralik, J. D.; Beck, P. D.; Laubach, M.; Chapin, J. K.; Kim, J.; Biggs, J.; Srinivasan, M. A.; Nicolelis, M. A. (2000): Real-time prediction of hand trajectory by ensembles of cortical neurons in primates. In: *Nature* 408(6810), S. 361–365
- Witt, K. (2013): Das Identitätsproblem der tiefen Hirnstimulation und einige seiner praktischen Implikationen. In: *Ethik in der Medizin* 25(1), S. 5–18
- Wolpaw, J. R. (2014): The BCI endeavor and the mission of this new journal. In: *Brain-Computer Interfaces* 1(1), S. 2–4

Anhang

Abbildungsverzeichnis 1.

Abb. II.1	Der Maschinenmensch »Maria« aus Metropolis (Replikat)	31
Abb. II.2	John D. Bernal (Foto nach 1964)	35
Abb. III.1	Schematische Darstellung eines ankommenden Aktionspotenzials	56
Abb. III.2	Illustration von ableitenden Verfahren	58
Abb. III.3	Visualisierung elektrischer Kontaktierung am peripheren Spinalnerv	66
Abb. III.4	Grundaufbau ableitender Anwendungen	69
Abb. III.5	Steuerung eines Roboterarms mittels implantierter Elektrode	76
Abb. III.6	Tiefe Hirnstimulation: Stimulation des Nucleus subthalamicus zur Behandlung von Parkinson	80
Abb. III.7	Aufbau eines Cochlea-Implantats	85
Abb. III.8	Ansatzpunkte für Sehprothesen	88
Abb. III.9	Neil Harbisson: Cyborg mit Eyeborg	92
Abb. III.10	Der Perception-Action-Learning-Loop	102
Abb. III.11	Mit schwenkbarem Laserscanner erzeugte 3-D-Karte eines Innenraums	111
Abb. III.12	Geschichtete Kontrollarchitektur eines Flugroboters	115
Abb. III.13	Der Roboter iCub	120
Abb. III.14	Leichtbauroboter LBR iiwa und Baxter	122
Abb. III.15	Packbot und LS3	126
Abb. III.16	Care-O-Bot 4 und Cosero	129
Abb. III.17	Paro und Pepper	132

Tabellenverzeichnis 2.

Tab. III.1	Kategorisierung neurotechnologischer Anwendungen	67
------------	--	----



**BÜRO FÜR TECHNIKFOLGEN-ABSCHÄTZUNG
BEIM DEUTSCHEN BUNDESTAG**

KARLSRUHER INSTITUT FÜR TECHNOLOGIE (KIT)

Neue Schönhauser Straße 10
10178 Berlin

Fon +49 30 28491-0
Fax +49 30 28491-119

buero@tab-beim-bundestag.de
www.tab-beim-bundestag.de