

# Künstliche Intelligenz und Neurotechnologie: Ethische Fragestellungen aus technologischen Innovationen

Elena Bleher\*

elena.bleher.eb@gmail.com

Ajla Glušac\*

aj.glusac@hotmail.de

Katharina Schnetzer\*

schnetzer.katharina@web.de

Lorena Frankenreiter\*

lorena.frankenreiter@hotmail.de

Veronika Öttl\*

oettl.veronika@gmail.com

Prof. Dr. Jörg Puchan\*

joerg.puchan@hm.edu

\*Hochschule München  
Fakultät Wirtschaftsingenieurwesen  
Lothstraße 64, 80335 München

## ABSTRACT

Fortschritte in den Bereichen Künstliche Intelligenz und Neurotechnologie haben das Potenzial, die Interaktion zwischen Mensch und Maschine neu zu definieren. So werden beispielsweise immer leistungsfähigere Roboterprothesen entwickelt, die in Verbindung mit dem menschlichen Nervensystem auch Anwendungen außerhalb therapeutischer Maßnahmen ermöglichen. Diese Entwicklungen wecken große Erwartungen in der Gesellschaft, werfen aber auch grundlegende ethische Fragen auf. Diese Arbeit untersucht, wie eine verantwortungsvolle Gestaltung von KI und Neurotechnologie auf gesellschaftspolitischer Ebene erreicht werden kann. Dabei steht insbesondere der ethische Diskurs im Mittelpunkt. Nach einem Überblick über die Grundlagen werden die Potenziale der Technologien dargestellt und die ethischen Herausforderungen skizziert. Abschließend werden mögliche Lösungsansätze aufgezeigt.

## SCHLÜSSELWÖRTER

Neurotechnologie, Künstliche Intelligenz, Ethik, Gesundheitswesen, Mensch-Maschine-Interaktion, Medizinische Forschung

## EINLEITUNG

*„I could have a Neuralink right now and you wouldn't know.“ (Musk 2020)*

Elon Musks kühne Aussage, er könne ein Implantat besitzen, das unbemerkt eine direkte Verbindung zwischen Gehirn und Computer herstellt, illustriert die vielversprechende Realität der Integration von menschlicher Intelligenz und Technologie. Die Verbindung von Gehirn und Computer ist längst keine Utopie mehr, sondern rückt mit zahlreichen Forschungsprojekten weltweit in greifbare Nähe. Ob zur Behandlung neurologischer und psychischer Erkrankungen oder zur Erforschung des menschlichen Gehirns - die Integration von Künstlicher Intelligenz und Neurotechnologie hat das Potenzial, die Neurowissenschaft zu revolutionieren. Im Zusammenhang mit dem Einsatz dieser Technologien ergeben sich jedoch auch eine Reihe von ethischen Fragen, die es zu beachten gilt. Eine sorgfältige Prüfung und Berücksichtigung ethischer Grundsätze bei der Anwendung dieser neurowissenschaftlichen Innovationen ist daher unerlässlich, um mögliche negative Auswirkungen auf die Gesellschaft sowie die individuellen Rechte und Freiheiten des Menschen zu vermeiden. Ziel der Arbeit ist es, einen Überblick über die Potenziale und ethischen Herausforderun-

gen der Integration von Künstlicher Intelligenz und Neurotechnologie zu geben und Lösungsansätze für eine verantwortungsvolle Entwicklung und Anwendung dieser Technologien aufzuzeigen.

Zu Beginn der Arbeit werden die zentralen Begriffe „Künstliche Intelligenz“ und „Neurotechnologie“ definiert und erläutert. Anschließend werden die Potenziale und zukünftigen Entwicklungen der Integration beider Technologien dargestellt. Zuletzt werden die damit einhergehenden ethischen Herausforderungen diskutiert und Lösungsansätze zur Vermeidung möglicher negativer Auswirkungen auf die Gesellschaft und den Menschen aufgezeigt. Die Arbeit schließt mit einem Fazit.

## KÜNSTLICHE INTELLIGENZ

Da es keine einheitliche Definition des Begriffs „Intelligenz“ gibt, existiert bislang auch keine allgemein anerkannte Definition für Künstliche Intelligenz (KI) (vgl. Bauer et al., S. 11). Eine Vielzahl der Definitionsansätze basiert jedoch auf dem Gedanken, dass KI ein Sammelbegriff für verschiedene Technologien ist, die es Computern ermöglichen, komplexe Aufgaben auszuführen, die normalerweise menschliche Intelligenz erfordern (vgl. Stuart Russell, 1 ff.).

Eines der wichtigsten Teilgebiete der KI ist das maschinelle Lernen (ML). Hierbei handelt es sich um den Einsatz von Algorithmen, die in der Lage sind, aus Daten Muster zu lernen und Aufgaben ohne explizite Programmierung zu bewältigen. Ein charakteristisches Merkmal

von ML ist, dass die Leistung durch gewonnene Erfahrung und Daten verbessert wird (vgl. Fraunhofer Institut 2023). Künstliche neuronale Netze (KNN) sind ein Untergebiet des MLs und von der biologischen Struktur des Gehirns inspiriert. Sie bestehen aus Knoten, die in Schichten angeordnet und durch gewichtete Verbindungen miteinander vernetzt sind. Dies ermöglicht die Verarbeitung von Eingabedaten zur Erstellung von Ausgabepronosen (vgl. Fraunhofer Institut 2023). Ein erfolgreiches Modell von KNNs sind die sogenannten tiefen neuronalen Netze (TNN). Dabei werden Informationen durch mehrere miteinander verbundene Schichten verarbeitet und in einer hierarchischen Struktur dargestellt. Dieses Modell kann dazu genutzt werden, um Daten unter Bedingungen zu klassifizieren, mit denen es zuvor nicht konfrontiert wurde. Neuronale Netze zu verwenden, um ML zu implementieren, steht in engem Zusammenhang mit den Neurowissenschaften (vgl. Ienca und Ignatiadis 2020, S. 77).

## NEUROTECHNOLOGIE

Die Neurotechnologie ist ein Forschungsgebiet, das durch die Integration von Technologien in die Neurowissenschaften entstanden ist. Die Neurowissenschaften befassen sich mit zwei Bereichen des Körpers, dem zentralen Nervensystem, zu dem das Gehirn und das Rückenmark gehören, und dem peripheren Nervensystem, das alle Nervenzellen außerhalb des zentralen Nervensystems umfasst. Alle diese Komponenten dienen der Aufnahme von Reizen wie Sehen, Tasten oder Schmecken und der anschließenden Weiterleitung und Verarbeitung der Informationen im Gehirn (vgl. Bear et al. 2018, 4 ff.). Einer der wichtigsten Bausteine des Nervensystems sind die Neuronen. Diese Nervenzellen sind in der Lage, über Verbindungsstellen, sogenannte Synapsen, Daten an andere Zellen weiterzuleiten und ermöglichen so die Kommunikation des gesamten Nervensystems mit dem Körper (vgl. Bear et al. 2018, S. 26). Diese Daten, auch Neurodaten genannt, können dank neuer Technologien und vor allem mit Hilfe von KI gesammelt, interpretiert und in manchen Fällen auch verändert werden.

Neurotechnologien können auf zwei unterschiedliche Arten Anwendung finden. Zum einen können die Daten direkt, invasiv über Implantate, oder indirekt, nicht-invasiv über tragbare Geräte auf der Hautoberfläche übertragen werden.

Darüber hinaus können die Neurotechnologien grob in die folgenden drei Bereiche eingeteilt werden:

- Neurosensorik: Neurodaten werden gelesen und analysiert,
- Neuromodulation: Neurodaten werden geschrieben und an das Gehirn übermittelt, und
- Kombinatorische/ bidirektionale Anwendungen: Die bidirektionale Integration von Neurosensorik und Neuromodulation führt zu einer wechselseitigen Interaktion, die nicht nur das Lesen und Analysieren von Neurodaten, sondern auch das Schreiben und Übermitteln von

Informationen an das Gehirn in einem nahtlosen und integrativen Prozess ermöglicht.

KI wird zunehmend in den Bereichen Neuromodulation und kombinatorische/bidirektionale Anwendungen eingesetzt, da sie die Interpretation neuronaler Daten und die Ableitung neuronaler Signale ermöglicht, wodurch ein gewisser Grad an Automatisierung und autonomer Steuerung erreicht wird (vgl. Berger und Rossi 2023, S. 61).

Abb. 1 zeigt die eben beschriebenen Bereiche mit Beispielen invasiver und non-invasiver Neurotechnologien.

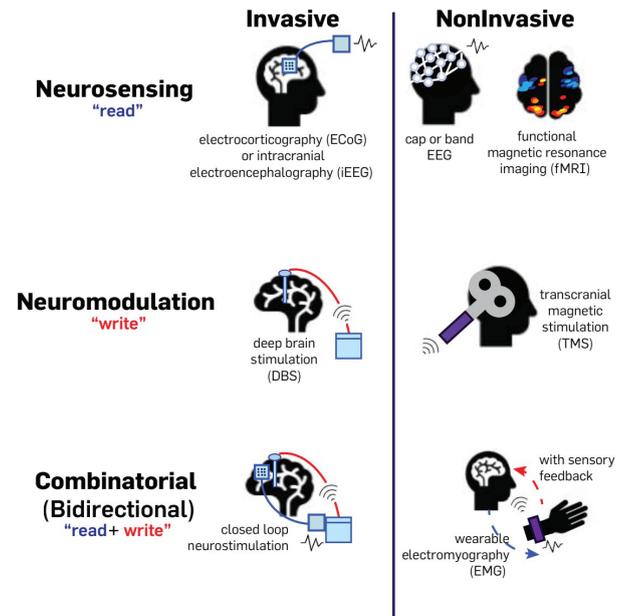


Abb. 1: Kategorisierung der Neurotechnologien (Berger und Rossi 2023, S. 62).

Bislang wurden Neurotechnologien vor allem im Gesundheitsbereich erforscht und eingesetzt, um bisher unheilbare Krankheiten zu behandeln und mehr über unser Gehirn zu erfahren. Inzwischen gibt es aber auch andere Anwendungsbereiche für Neurotechnologien. Einer davon ist die Gaming-Industrie, einer davon ist die Gaming-Industrie, welche Elektroenzephalographie-Headsets (EEG-Headsets) einsetzt, um Videospiele per Gedanken zu steuern (vgl. Paek et al. 2021, S. 11). Die Potenziale der Neurotechnologien sind vielfältig und werden nachfolgend beschrieben.

Allerdings steigt aufgrund der rasanten Entwicklung der Neurotechnologien und der Tatsache, dass wir dadurch direkt als auch indirekt in unserem Denken und Entscheiden beeinflusst werden können, die Skepsis gegenüber möglichen negativen Auswirkungen dieser Entwicklungen. Um die ethischen Fragen in diesem Bereich zu klären, wurde die Neuroethik als Forschungsgebiet geschaffen (vgl. Robinson et al. 2022, S. 2057).

## POTENZIALE UND ENTWICKLUNGEN

Die Verbindung von KI und Neurotechnologie bietet vielversprechende Zukunftsaussichten für die Neurowissenschaften und die Gesellschaft im Allgemeinen. In diesem Kontext werden im Folgenden aktuelle Beiträge der

KI-gestützten Neurotechnologie im Gesundheitswesen, in der Grundlagenforschung und in alltäglichen Prozessen beleuchtet und ein Ausblick auf zukünftige Entwicklungen gegeben.

## **POTENZIALE IN DER DIAGNOSTIK**

Bereits seit den 1970er Jahren wird das Gesundheitswesen als ein klassisches Anwendungsgebiet der KI betrachtet (vgl. Patel et al. 2009, S. 6). Insbesondere im Bereich neurologischer Erkrankungen zeichnen sich aufgrund des exponentiellen Wachstums von KI-Technologien bedeutende Potenziale zur Verbesserung von Prognosen sowie Diagnose- und Therapiemaßnahmen ab (vgl. Topol und Verghese 2019, 25 ff.).

Durch den Einsatz von ML-Modellen eröffnen sich vielfältige Möglichkeiten, komplexe Zusammenhänge zwischen neuronalen Aktivitätsmustern und neurologischen Störungen zu identifizieren. Dies führt zu einer Verbesserung der Vorhersage schwerwiegender Erkrankungen wie Epilepsie, Hirntumoren und Alzheimer. Im Bereich der psychischen Gesundheit können sie zur Untersuchung von Entwicklungsstörungen wie Autismus und zur diagnostischen Prognose in neuropsychiatrischen Bereichen wie Psychosen eingesetzt werden (vgl. Ienca und Ignatiadis 2020, 78 f.). Darüber hinaus können subtile Bildmerkmale erkannt werden, die von konventionellen Systemen aufgrund von Bildfehlern oder Unterschieden in der Anatomie möglicherweise übersehen oder falsch interpretiert werden. Die gewonnenen Informationen können anschließend genutzt werden, um Muster und Beziehungen zwischen den Daten zu identifizieren. Auf diese Weise können strukturelle Veränderungen im Gehirn, die potenziell auf einen Hirntumor hindeuten, frühzeitig erkannt werden (vgl. Ienca und Ignatiadis 2020, S. 80).

Die Anwendung von KI in den klinischen Neurowissenschaften zielt jedoch nicht nur auf algorithmische Vorhersagen und Diagnosen ab, sondern ermöglicht auch die Optimierung von Neurotechnologien. Dies eröffnet erfolgsversprechende Behandlungsoptionen für eine Vielzahl neurologischer und psychischer Erkrankungen (vgl. Vukelic 2023).

## **VEREINFACHUNG VON PERSONALISIERTEN THERAPIEANSÄTZEN**

Im Falle einer neurologischen Erkrankung kann die Kombination von KI und invasiver Neurotechnologie dabei helfen, die Aktivitäten des Gehirns zu überwachen und gegebenenfalls zu regulieren und zu steuern. Hierbei werden elektrische Impulse an die betroffenen Bereiche im Gehirn abgegeben, wodurch die Symptome reduziert und die Lebensqualität der Patienten gesteigert werden kann. Krankheiten wie Epilepsie und Parkinson können dadurch effektiv behandelt werden (vgl. Müller und Rötter, 36 ff.).

Ein weiteres Beispiel, wie durch die Nutzung von KI und Neurotechnologie die Lebensqualität der einzelnen Patienten verbessert werden kann, ist die Integration in die neurorehabilitative Therapie. Auch hier wird das Gehirn

durch die Kombination von Gehirn-Computer-Schnittstellen (BCIs) mit KI-Algorithmen überwacht, Aktivitäten erfasst und fallbezogen gesteuert. Dadurch ist es zum Beispiel möglich neurologische Beeinträchtigungen, hervorgerufen bspw. durch einen Schlaganfall oder Rückenmarksverletzungen, zu behandeln. Die Steuerung von Prothesen wird dadurch für den Patienten vereinfacht und die Wiederherstellung von motorischen Fähigkeiten unterstützt (vgl. Kehl, 151 ff.).

Durch KI und invasive Neurotechnologie ergeben sich weiterhin neue Chancen im Bereich der personalisierten Medizin. Bei der Erstellung von Behandlungsplänen kann die KI Informationen aus großen Datenbanken abrufen und mit den gesammelten persönlichen Daten des Gehirns abgleichen. Dadurch ist es möglich, maßgeschneiderte Therapien auf die Bedürfnisse und Eigenschaften jedes einzelnen Patienten abzustimmen und den Fortschritt der Behandlung optimal zu überwachen (vgl. Dilsizian und Siegel 2014; vgl. Iliashenko et al. 2019).

## **WEGBEREITER IN DER MEDIZINISCHEN FORSCHUNG**

Neben den bereits erwähnten Möglichkeiten, die sich aus der Kombination von KI und Neurotechnologie in der angewandten Medizin ergeben, bieten sie auch vielversprechende Chancen in der Forschung.

Ein möglicher aussichtsvoller Aspekt ist der effizientere Einsatz von Ressourcen. Repräsentative Studien fundieren meist auf umfangreichen Datensätzen, welche mithilfe zeitintensiver Analysen und Auswertungen bearbeitet werden. Durch die effektive Integration von KI können diese aufwendigen Prozesse optimiert werden. Somit wäre es möglich, die menschlichen Arbeitskräfte, welche oftmals kostenintensiv sind, an anderer bedeutenderer Stelle einzusetzen. Zusätzlich ist die KI bspw. in der Lage, mittels der Informationen bestimmte Hirnaktivitätsmuster zu erlernen und sich durch immer neue Daten zu verbessern. Aufgrund dessen wäre es Forschern möglich, Thesen virtuell zu testen, bevor hochpreisige Investitionen und Testreihen gestartet werden müssen (vgl. Savage 2019, 15 ff.).

Der Lerneffekt und die Simulation von Hirnströmen können zudem dabei hilfreich sein, die Bildgebung des Gehirns zu verbessern. Klassische bildgebende Verfahren wie die Magnetresonanztomographie (MRT) können durch diverse Störfaktoren oftmals keine optimale Bildqualität erzielen. KI-Algorithmen können an dieser Stelle dazu verwendet werden, fehlende bzw. unklare Bereiche zu rekonstruieren. Damit können präzisere, detailliertere und umfassendere Bilder des Gehirns erstellt werden, welche dazu beitragen können, neue Erkenntnisse über neurologische Zusammenhänge zu erlangen. Mithilfe dieses Wissens können neue Ansätze für Diagnosen und Behandlungen von Krankheiten wie Epilepsie oder Schizophrenie entwickelt werden (vgl. Lubkowitz 2022, 44 f.; vgl. Lorenz 2022, S. 14).

## REVOLUTIONÄRE CHANCEN AUßERHALB DER MEDIZIN

Ein Leben ohne Interaktion mit elektronischen Geräten ist in weiten Teilen der Bevölkerung kaum mehr möglich (vgl. BMWK 2023). Durch die Kopplung aus bidirektionalen Neurotechnologien in Kombination mit KI können neue Chancen im Alltag auf dem Gebiet der Digitalisierung geschaffen werden. Wie bereits erwähnt, finden sich in der Gaming-Industrie unter anderem neue Potenziale in Bezug auf das Spielerlebnis. Durch die Verarbeitung von Gehirnsignalen können Spieler dann bspw. ihre Charaktere nur durch ihre Gedanken steuern, anstatt physische Eingabegeräte zu verwenden. Das würde unter anderem die Immersion und Interaktion in Videospielen intensivieren (vgl. Paszkiel et al. 2021, 109 ff.). Des Weiteren ergeben sich Potenziale für die Neugestaltung von Arbeitsplätzen. Gedanken können direkt in digitale Dokumente überführt und so die Produktivität gesteigert werden. Bereichsübergreifend wird die Zusammenführung von KI und Neurotechnologien dazu beitragen, die Barrierefreiheit im Alltag auszubauen. Menschen mit eingeschränkter Mobilität oder anderen Beeinträchtigungen wird der Zugang zu bisher oft unzugänglichen Bereichen ermöglicht und damit die Inklusion gestärkt (vgl. Vukelic 2023).

Studien haben herausgefunden, dass diverse psychische Störungen, wie z.B. Pädophilie und Psychopathie, anhand von Auffälligkeiten bei den neuronalen Aktivitäten erkannt werden können. Bestimmte Bereiche des Gehirns werden hier in einer anderen, abnormalen Intensität stimuliert (vgl. Ponseti et al. 2012, 187 ff.; vgl. Neurologie: Schalten die Gehirne Pädophiler einfach falsch? 2007). Im Rahmen der Strafverfolgung ist es beispielsweise möglich, durch gezielte Impulse an das Gehirn "illegale" Gedanken zu unterdrücken, um so strafbare Handlungen zu vermeiden. Dadurch wird das Allgemeinwohl der Menschheit verbessert und die Kriminalitätsrate gesenkt. Trotz der vielversprechenden Potenziale und Vorteile KI-gestützter Neurotechnologien gibt es auch wachsende Bedenken und kontroverse Diskussionen über die potenziell destruktiven Auswirkungen dieser Technologie.

## ETHISCHER DISKURS

Das folgende Kapitel befasst sich mit ethischen Konflikten und Fragen, die bei der Integration von KI und Neurotechnologien auftreten können. Anschließend werden mögliche Lösungsansätze vorgestellt.

## FRAGESTELLUNGEN

Die unabhängige *Hochrangige Expertengruppe für Künstliche Intelligenz* (HLEG AI) veröffentlichte 2019 Ethik-Leitlinien für eine vertrauenswürdige KI. Die Entwicklung und der Einsatz solcher Leitlinien begründen sich in ethischen Bedenken bezüglich digitaler Spaltungen, Diskriminierungen, bis hin zu ernsthaften Gesundheits- und Sicherheitsproblemen durch den Einsatz von KI (vgl. Stahl et al. 2022, S. 1). Viele dieser ethischen

Fragestellungen finden sich auch innerhalb der Neuroethik. Im Folgenden werden anhand der in Abb. 2 dargestellten Leitlinien einige dieser ethischen Fragestellungen aufgefasst und diskutiert.

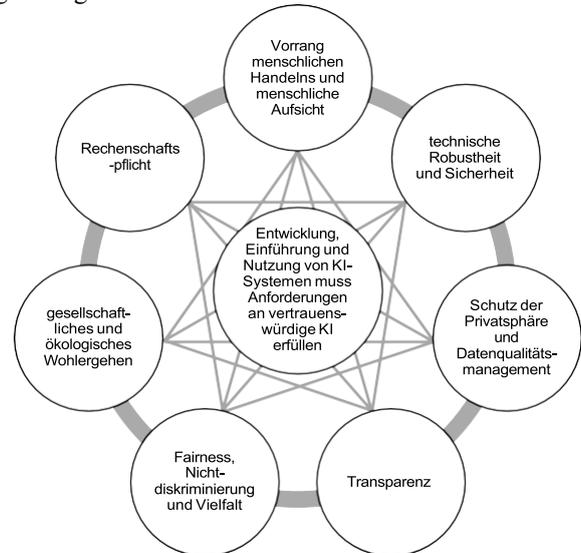


Abb. 2: Ethische Herausforderungen durch KI-gestützte Neurotechnologien (*Selbsterstellte Darstellung in Anlehnung an Unabhängige hochrangige Expertengruppe für Künstliche Intelligenz 2018, S. 18*).

*Vorrang menschlichen Handelns und menschliche Aufsicht*

Wird der Mensch zukünftig von KIs kontrolliert?

Durch den Einsatz KI-gestützter Neurotechnologien werden Maschinen, Körper und Psyche zukünftig zunehmend miteinander verschmelzen. Durch den Einsatz in der Medizin können darüber hinaus Gehirnaktivitäten überwacht und gesteuert werden. Hierbei konnte beobachtet werden, dass KI-Komponenten unter bestimmten Umständen die Entscheidungsfähigkeit der Patienten überlagern. Dies führt dazu, dass die Handlungsfähigkeit, der subjektive Wille und somit die Kontrolle der Patienten teilweise an ein autonomes System abgegeben wird (vgl. Ienca und Ignatiadis 2020, S. 84).

Der Einfluss auf das Entscheidungsverhalten und Handlungsvermögen der Menschen kann teilweise weder gemessen, noch kontrolliert werden. Somit wird nicht nur die Autonomie des Menschen gefährdet, sondern auch seine Identität. Diese wird bspw. durch Kultur, Emotionen, Erinnerungen oder Beziehungen geformt und kann durch das direkte Eingreifen der Neurotechnologie beeinflusst oder verändert werden (vgl. DIN e. V. und DKE Deutsche Kommission Elektrotechnik, S. 38).

Zusätzlich besteht die Gefahr, dass die KI durch die Integration von Neurotechnologien und der Fähigkeit zur Simulation und Reproduktion menschlichen Denkens und Verhaltens noch größere Autonomie erlangt. Dadurch entsteht das so genannte „Wall-E-Problem“, wonach die Menschheit durch die vollständige Abhängigkeit von KI-Systemen ihre Fähigkeit zur Selbstversorgung verliert und auf eine passive Empfängerrolle reduziert wird. Kulturelle und intellektuelle Vielfalt gehen somit verloren, was langfristig zur Stagnation gesellschaftlicher Entwicklung führt (vgl. Doya et al. 2022, S. 547).

### *Technische Robustheit und Sicherheit*

#### Wie zuverlässig und sicher sind KI und Neurotechnologie?

Ein wichtiger technischer Aspekt bei der ethischen Betrachtung von Neurotechnologien ist die Zuverlässigkeit der eingesetzten Systeme. Da KI und ML auf Statistiken beruhen, bleibt selbst bei mit umfangreichen Datensätzen trainierten und starken Systemen ein gewisser Prozentsatz an möglichen Fehlentscheidungen (vgl. Berger und Rossi 2023, S. 60). Je mehr Menschen sich auf eine KI verlassen und je folgenschwerer die Entscheidungen sind, desto mehr Gefahr verbirgt sich in auftretenden Fehlern (vgl. Santosh 2023, S. 10; vgl. Barton und Pöppelbuß 2022, S. 473).

Ein weiteres erkanntes Risiko bildet das Einfangen, Überwachen sowie Steuern von Gehirnströmen. Während diese Art der Neurotechnologie große Fortschritte in der Behandlung verschiedener neurologischer Krankheiten verspricht, werden Bedenken zur Cybersicherheit der Patienten geäußert. Jeder sich im Netz befindende Computer bildet ein mögliches Ziel für eine Cyber Attacke. Wenn also Computer durch Neurotechnologie mit Gehirnen verbunden sind, besteht das potenzielle Risiko einer „Übernahme“ des Gehirns durch Dritte (vgl. Doya et al. 2022, S. 547).

#### Schutz der Privatsphäre und Datenqualitätsmanagement Inwieweit ist die Privatsphäre des Menschen im Zeitalter der KI-Neurotechnologie gewährleistet?

Wie bereits erwähnt, kann KI in Verbindung mit Neurotechnologie bestimmte Krankheitsbilder erkennen, noch bevor klinische Symptome auftreten. Ebenso lassen sich gewisse Entscheidungen oder persönliche Präferenzen bestimmen. Diese gewonnenen Daten müssen durch die hohe Individualität und Sensibilität besonders geschützt werden und unterliegen daher dem gesonderten Feld der „Neuroprivacy“. Wie ein Gehirn auf Situationen, Reize etc. reagiert, ist hochindividuell und die gesammelten Daten daher sehr intim (vgl. Unabhängige hochrangige Expertengruppe für Künstliche Intelligenz 2018, S. 21). Werden diese sehr privaten Gedanken, Einstellungen und Werte für jeden zugänglich, kann dies unter anderem zu Diskriminierung führen. Aufgrund der zum Teil sehr persönlichen Daten ist eine Anonymisierung der Informationen, die eine genaue Identifizierung von Personen verhindert, nur bedingt möglich (vgl. da Silva Castanheira et al. 2021).

Eine große Gefahr liegt hierbei im „Überwachungskapitalismus“, also der Kommerzialisierung der persönlichen Daten im Zuge von Profitsteigerungen (vgl. Stahl et al. 2022, S. 39). So können veröffentlichte Daten, die den Gesundheitszustand einer Person beschreiben oder vorhersagen, von Arbeitgebern oder Krankenversicherungen gegen die Interessen dieser Person verwendet werden (vgl. Ienca und Ignatiadis 2020, S. 84; vgl. Stahl et al. 2022, S. 39).

Eine zusätzliche Herausforderung lässt sich im Rahmen des Datenqualitätsmanagements identifizieren. Unpräzise Datenerhebung und -auswertung sind z.B. auf unzu-

reichende Kenntnisse des Nervensystems zurückzuführen (vgl. Berger und Rossi 2023, S. 66). Weitere Fehlerquellen können durch unausgereifte Sensoren, Methoden und Techniken in der Anwendung der Neurotechnologie sowie durch mangelnde Expertise in der Auswertung der so gewonnenen Daten entstehen. Die Kombination dieser Fakten führt zu Ungenauigkeiten in den Analysen. Mangelnde Präzision und daraus resultierende mögliche Fehlschlüsse führen zu hohen Unsicherheitsfaktoren und damit zu einer geringen Robustheit der Technologie (vgl. Fairclough und Lotte 2020, S. 2). Außerdem wird derzeit nur ein kleiner Teil der gesamten Gehirnaktivität verstanden und aktiv genutzt. Dies führt z. B. zu Verzerrungen in den Datensätzen für das Training der KI, zu weiteren Fehlschlüssen der KI und in der Folge zu Fehlschlüssen oder Fehleinschätzungen in Bezug auf den Patienten und seinen geistigen oder körperlichen Gesundheitszustand (vgl. Barton und Pöppelbuß 2022, S. 473). Eine daraus resultierende Fehldiagnose kann weitreichende Folgen für die Betroffenen haben. Erhobene Daten und daraus folgende Auswertungen sind daher einer strengen Kontrolle und damit einem strengen Qualitätsmanagement zu unterziehen.

### *Transparenz*

#### Sind KI-gestützte Neurotechnologien nachvollziehbar gestaltet?

Die Transparenz KI-gestützter Neurotechnologien gilt als Schlüsselfaktor für die Akzeptanz und das Vertrauen der Nutzenden (vgl. Larsson und Heintz 2020). Je komplexer Neurotechnologien wie z. B. BCIs sind, desto schwieriger ist es, die zugrunde liegenden Prozesse zu verstehen (vgl. Santosh 2023, S. 21). Durch den Einsatz von KI-Algorithmen zur Umsetzung neuronaler Signale in motorische Aktionen wird zusätzlich eine „Black Box“ in den ohnehin komplexen Vorgang der Transformation neuronaler Inputs in aktionsfähige Outputs eingeführt (vgl. Rainey und Erden 2020, 2444 ff.). Dies hat zur Folge, dass selbst für Entwickler und Anwender die genauen Prozesse der neuronalen Signalverarbeitung oft undurchsichtig und teilweise nicht erklärbar sind.

### *Fairness, Nichtdiskriminierung und Vielfalt*

#### Inwieweit werden Fairness, Nichtdiskriminierung und Vielfalt gefördert?

Eine gerechte Verteilung von Ressourcen und Chancen ist eine ethische Grundlage für eine gleichberechtigte und demokratische Gesellschaft. Der Erwerb und die Nutzung dieser Technologien erfordern oft finanzielle Mittel und spezifische Fachkenntnisse, die nicht jedermann zur Verfügung stehen. Die mögliche Folge ist eine Benachteiligung einkommensschwacher Gruppen, insbesondere in Entwicklungsländern. Sie können somit von den Vorteilen dieser Technologien ausgeschlossen werden. Dadurch wird die bereits bestehende Marginalisierung benachteiligter Gruppen verstärkt und sozioökonomische Disparitäten verschärft (vgl. Friedrich 2021, S. 129).

Ein weiteres bedeutendes Problem ist die potenzielle Diskriminierung infolge algorithmischer Verzerrungen. Derartige systematische Fehler können entstehen, wenn die

zum Trainieren der KI-Algorithmen verwendeten Datensätze nicht ausreichend repräsentativ sind und bestimmte Gruppen vernachlässigt werden. Die daraus resultierende algorithmische Diskriminierung kann insbesondere bei unterrepräsentierten Gruppen wie ethnischen Minderheiten oder Patienten mit seltenen neurologischen Erkrankungen zu ungenauen Vorhersagen und Behandlungsstandards in der Neurotechnologie führen (vgl. Ienca und Ignatiadis 2020, S. 83).

#### *Gesellschaftliches und ökologisches Wohlergehen*

##### Ist der Einsatz von KI und Neurotechnologien sozialverträglich und ökologisch nachhaltig?

Die Konvergenz von KI und Neurotechnologien bietet ein breites Spektrum an Möglichkeiten in Medizin, Forschung und Alltag. Zugleich erfordert ihre Integration eine reflektierte Debatte über ihre ethischen Implikationen hinsichtlich sozialer Verträglichkeit und ökologischer Nachhaltigkeit. Ein wesentlicher Aspekt gesellschaftlichen Wohlergehens sind kontinuierliches Wirtschaftswachstum und eine breite Wohlstandsverteilung (vgl. McKinsey Global Institute 2018, 2 ff.). Mit den aufstrebenden Möglichkeiten der KI in Verbindung mit Neurotechnologie eröffnen sich innovative Perspektiven, die Leistungsfähigkeit zu verbessern und gleichzeitig das Wohlbefinden der Menschen zu fördern (vgl. Farahany 2023). Allerdings birgt die zunehmende Automatisierung von Arbeitsprozessen das Risiko, bestimmte manuelle Tätigkeiten durch diese Technologien zu ersetzen. Dadurch kann es zu einem Verlust von Arbeitsplätzen und in der Folge zu einer Verschärfung von Ungleichheiten auf dem Arbeitsmarkt kommen. Die Folge sind enorme Disparitäten innerhalb und zwischen den Ländern, die wiederum soziale und internationale Spannungen hervorrufen (vgl. PwC 2023).

Der Einfluss von KI und Neurotechnologien auf das gesellschaftliche Wohlergehen darf nicht isoliert betrachtet werden, sondern erfordert die Berücksichtigung ökologischer Aspekte. In einigen Publikationen wird bspw. die Befürchtung geäußert, dass durch den Einsatz von KI der Energiebedarf und damit die CO<sub>2</sub>-Emissionen steigen werden (vgl. Thompson et al. 2021, S. 52; vgl. Ekin 2023).

#### *Rechenschaftspflicht*

##### Wer haftet bei Schäden durch neurotechnologische KI-Anwendungen?

Wie bereits erwähnt, resultieren durch die Kombination aus KI und Neurotechnologien diverse Unsicherheiten in Bezug auf die Zuverlässigkeit und Fehlerfreiheit. Bei der Anwendung dieser Technologien dürfen demnach die potenziellen Risiken und Schäden nicht außer Acht gelassen werden. KI-bedingte Fehlfunktionen können schwerwiegende Auswirkungen auf Menschen und Umwelt haben. Bei rechtlich und ethisch fragwürdigen Handlungen ist es schwierig, den Verantwortlichen zu identifizieren. Es ist nicht nachvollziehbar, ob es sich um eine vorsätzliche Handlung oder um eine Fehlfunktion der KI handelt (vgl. Davidoff 2020).

Zusammenfassend lässt sich sagen, dass die sieben ethischen Leitlinien im Zusammenhang mit KI und Neurotechnologien sowohl Aufmerksamkeit als auch entsprechende Handlungen erfordern.

## **LÖSUNGSANSÄTZE**

Um die ethischen Herausforderungen im Umgang mit KI-gestützten Neurotechnologien zu bewältigen, wurden im Rahmen der Ausarbeitung diverse Lösungsansätze identifiziert, welche nachfolgend dargestellt werden.

Zur Bewältigung der genannten Herausforderungen sind präventive Maßnahmen erforderlich. Ein möglicher Ansatz ist, zusätzlich zu den existierenden Menschenrechten einen neuroethischen rechtlichen Rahmen zu schaffen. Dadurch wird das Recht auf Schutz der gedanklichen Privatsphäre vor unerwünschter Aufzeichnung und Manipulation sichergestellt. In diesem Kontext müssen entsprechende Rechtsnormen geschaffen, bzw. weiterentwickelt werden. Jedem Individuum wird damit das Recht eingeräumt, selbstbestimmt über die Speicherung und Freigabe seiner Neurodaten, sowie über den Umfang ihrer Nutzung zu entscheiden (vgl. Hertz 2023, 2 ff.). Alle Entitäten, die neuronale Daten erheben, analysieren, nutzen und weitergeben, müssen diese Rechte des Einzelnen in Bezug auf seine Daten anerkennen.

Für die praktische Umsetzung dieser Rechte müssen Standards festgelegt werden. Diese erfordern eine ausdrückliche Zustimmung zur Erhebung von Neurodaten und sehen spezifische Einwilligungen hinsichtlich der Verwendung, des Zwecks und der Dauer der Datennutzung vor (vgl. Goering et al. 2021, S. 378). Darüber hinaus ist es wichtig zu berücksichtigen, inwiefern Menschen, selbst wenn sie eigenverantwortlich über die Speicherung und Nutzung ihrer Neurodaten entscheiden dürfen, die Tragweite dieser Entscheidung begreifen können. Um diese Wissenslücke zu schließen, sind möglicherweise zusätzliche Maßnahmen erforderlich. Um unbefugten Zugriff auf die Daten zu verhindern, wird die Verwendung einer Verschlüsselung des gesamten Datentransfers vom Ort der Aufzeichnung im Gehirn bis zum neuronalen Schnittstellengerät empfohlen. Hierbei können moderne Technologien wie die „homomorphe Verschlüsselung“ eingesetzt werden, um eine sichere Analyse der Daten zu ermöglichen (vgl. Goering et al. 2021, S. 378).

Die genannten Maßnahmen zur sicheren Handhabung und Verarbeitung neuronaler Daten können dazu beitragen, das Vertrauen der Gesellschaft in KI und Neurotechnologien zu stärken. Zudem ist es wichtig von Beginn des Prozesses an bis zur Nutzungsphase, Transparenz für alle Beteiligten, d.h. sowohl für Forschungsteams oder Unternehmen als auch für die Gesellschaft, zu schaffen. Dadurch wird das Vertrauen weiter gestärkt. Hierzu können verschiedene Kommunikationswege zur Ansprache von Verbrauchern verwendet werden (vgl. Goering et al. 2021, S. 379).

Um jedoch sicherzustellen, dass alle Bevölkerungsgruppen gleichermaßen von diesen Technologien profitieren, ist ein gerechter Zugang von großer Bedeutung. Regierungen können hierbei eine wichtige Rolle spielen, indem

sie Verteilungs-, Subventions- und Anreizprogramme schaffen (vgl. Gaudry et al. 2021). Durch eine bessere Zugänglichkeit dieser Technologien können sie von einem größeren Nutzerkreis in Anspruch genommen werden. Dies führt zu einer Vielfalt von Perspektiven und Erfahrungen und kann dazu beitragen, die für die Entwicklung von KI-Modellen verwendeten Datensätze zu diversifizieren. Dadurch wird die Anpassungsfähigkeit der KI-Modelle an unterschiedliche Bedürfnisse und Realitäten verbessert und ihre Genauigkeit und Zuverlässigkeit erhöht.

Um beispielsweise Arbeitsplatzverluste durch KI zu vermeiden, sollten Arbeitnehmer die Möglichkeit zur kontinuierlichen Weiterbildung erhalten und die Konzentration auf kreative und schwer automatisierbare Tätigkeiten gefördert werden. Unternehmen sollten außerdem darauf achten, auf einen offenen sozialen Dialog zu setzen (vgl. Hirsch-Kreinsen und Karacic 2019, 112 ff.).

Klare Haftungsregelungen, verstärkte Überwachungssysteme zur Früherkennung von Risiken und die Einführung ethischer Richtlinien für KI-Entwickler können dazu beitragen, die rechenschaftspflichtige Anwendung von KI-gestützter Neurotechnologie zu gewährleisten. Transparente Mechanismen erleichtern die Identifikation von Verantwortlichen bei Schäden und fördern eine verantwortungsbewusste Nutzung dieser Technologien (vgl. Europäische Kommission 2022).

Verallgemeinernd lässt sich festhalten, dass die rasante Entwicklung der KI-gestützten Neurotechnologien weiteren Forschungsbedarf aufwirft. Dabei gilt es, ethische Konflikte frühzeitig zu identifizieren und die Erkenntnisse in die Entwicklung und Verbesserung der Technologien einfließen zu lassen. Weiterhin müssen erkannte ethische Fragestellungen durch die stetige Weiterentwicklung kontinuierlich evaluiert werden (vgl. Unabhängige hochrangige Expertengruppe für Künstliche Intelligenz 2018, S. 45). Hierbei ist es wichtig, nicht isoliert auf lokaler Ebene Richtlinien und Standards zu schaffen, sondern internationale Regelungen zu entwickeln und umzusetzen. So kann in Zukunft eine global sichere, effiziente und vertrauenswürdige KI-gestützte Neurotechnologie gewährleistet werden.

## FAZIT

Ziel der Arbeit war es, einen Überblick über die Potenziale und ethischen Herausforderungen der Integration von KI und Neurotechnologie zu geben und Lösungsansätze für eine verantwortungsvolle Entwicklung und Anwendung dieser Technologien aufzuzeigen.

KI-gestützte Neurotechnologien bieten zahlreiche Möglichkeiten, Diagnoseverfahren, personalisierte Therapien und die medizinische Forschung wirksamer und präziser zu gestalten. Auch im Alltag können sie bspw. dazu beitragen, neue Arbeitsplätze zu schaffen und die Barrierefreiheit zu verbessern.

Die vorliegende Arbeit verdeutlicht aber auch, die erheblichen ethischen Implikationen, die mit der Konvergenz von KI und Neurotechnologie einhergehen. Insbesondere die Aspekte der technischen Robustheit und der Wahrung der Privatsphäre der Nutzenden stellen fundamentale

ethische Herausforderungen dar, die es zu adressieren gilt. Zudem ist es unabdingbar, Transparenz und Nichtdiskriminierung als grundlegende Prinzipien bei der Entwicklung und Anwendung dieser Technologien zu berücksichtigen.

In Zukunft wird es daher von großer Bedeutung sein, dass Wissenschaft und Gesellschaft gemeinsam an der Entwicklung und Implementierung internationaler Standards und Richtlinien arbeiten.

## LITERATURVERZEICHNIS

Barton, Marie-Christin; Pöppelbuß, Jens (2022): Prinzipien für die ethische Nutzung künstlicher Intelligenz. In: *HMD* 59 (2), S. 468–481. DOI: 10.1365/s40702-022-00850-3.

Bauer, Wilhelm; Ganz, Walter; Hämmerle, Moritz; Renner, Thomas: *KÜNSTLICHE INTELLIGENZ IN DER UNTERNEHMENSPRAXIS*. Online verfügbar unter <https://www.smart-ai-work.de/wp-content/uploads/2020/01/kuenstliche-intelligenz-in-der-unternehmenspraxis.pdf>.

Bear, Mark F.; Connors, Barry W.; Paradiso, Michael A. (2018): *Neurowissenschaften. Ein grundlegendes Lehrbuch für Biologie, Medizin und Psychologie*. 4. Aufl. 2018. Hg. v. Andreas K. Engel. Berlin, Heidelberg: Springer Berlin Heidelberg. Online verfügbar unter <http://nbn-resolving.org/urn:nbn:de:bsz:31-epflicht-1555094>.

Berger, Sara; Rossi, Francesca (2023): AI and Neurotechnology. In: *Commun. ACM* 66 (3), S. 58–68. DOI: 10.1145/3529088.

BMWK (2023): *Den digitalen Wandel gestalten*. BMWI. Online verfügbar unter <https://www.bmwk.de/Redaktion/DE/Dossier/digitalisierung.html>, zuletzt aktualisiert am 04.05.2023, zuletzt geprüft am 04.05.2023.

da Silva Castanheira, Jason; Orozco Perez, Hector Domingo; Misic, Bratislav; Baillet, Sylvain (2021): Brief segments of neurophysiological activity enable individual differentiation. In: *Nature communications* 12 (1), S. 5713. DOI: 10.1038/s41467-021-25895-8.

Davidoff, Erika J. (2020): Agency and Accountability: Ethical Considerations for Brain-Computer Interfaces. In: *The Rutgers journal of bioethics* 11, S. 9–20. Online verfügbar unter <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7654969/>.

Dilsizian, Steven E.; Siegel, Eliot L. (2014): Artificial intelligence in medicine and cardiac imaging: harnessing big data and advanced computing to provide personalized medical diagnosis and treatment. In: *Current cardiology reports* 16 (1), S. 441. DOI: 10.1007/s11886-013-0441-8.

DIN e. V.; DKE Deutsche Kommission Elektrotechnik: Ethik und Künstliche Intelligenz. Was können technische Normen und Standards leisten? Online verfügbar unter <https://www.din.de/resource/blob/754724/00dcbccc21399e13872b2b6120369e74/whitepaper-ki-ethikaspekte-data.pdf>.

Doya, Kenji; Ema, Arisa; Kitano, Hiroaki; Sakagami, Masamichi; Russell, Stuart (2022): Social impact and governance of AI and neurotechnologies. In: *Neural networks: the official journal of the International Neural Network Society* 152, S. 542–554. DOI: 10.1016/j.neunet.2022.05.012.

Ekin, Annette (2023): AI can help us fight climate change. But it has an energy problem, too. Online verfügbar unter <https://ec.europa.eu/research-and-innovation/en/horizon-magazine/ai-can-help-us-fight-climate-change-it-has-energy-problem-too>, zuletzt aktualisiert am 01.05.2023, zuletzt geprüft am 01.05.2023.

Europäische Kommission (2022): RICHTLINIE DES EUROPÄISCHEN PARLAMENTS UND DES RATES zur Anpassung der Vorschriften über außervertragliche zivilrechtliche Haftung an künstliche Intelligenz (Richtlinie über KI-Haftung). Online verfügbar unter <https://eur-lex.europa.eu/legal-content/DE/TXT/PDF/?uri=CELEX:52022PC0496>.

Fairclough, Stephen H.; Lotte, Fabien (2020): Grand Challenges in Neurotechnology and System Neuroergonomics. In: *Front. Neuroergonomics* 1, Artikel 602504. DOI: 10.3389/fnrgo.2020.602504.

Farahany, Nita A. (2023): Neurotech at Work. Online verfügbar unter <https://hbr.org/2023/03/neurotech-at-work>, zuletzt aktualisiert am 14.03.2023, zuletzt geprüft am 30.04.2023.

Fraunhofer Institut (2023): Künstliche Intelligenz (KI) und maschinelles Lernen - Fraunhofer IKS. Online verfügbar unter <https://www.iks.fraunhofer.de/de/themen/kuenstliche-intelligenz.html>, zuletzt aktualisiert am 05.05.2023, zuletzt geprüft am 05.05.2023.

Friedrich, Orsolya (2021): Clinical Neurotechnology Meets Artificial Intelligence. Philosophical, Ethical, Legal and Social Implications. Unter Mitarbeit von Andreas Wolkenstein, Christoph Bublitz, Ralf J. Jox und Eric Racine. Cham: Springer International Publishing AG (Advances in Neuroethics Ser). Online verfügbar unter <https://ebookcentral.proquest.com/lib/kxp/detail.action?docID=6511455>.

Gaudry, Kate S.; Ayaz, Hasan; Bedows, Avery; Celnik, Pablo; Eagleman, David; Grover, Pulkit et al. (2021): Projections and the Potential Societal Impact of the Future of Neurotechnologies. In: *Frontiers in neuroscience* 15, S. 658930. DOI: 10.3389/fnins.2021.658930.

Goering, Sara; Klein, Eran; Specker Sullivan, Laura; Wexler, Anna; Agüera Y Arcas, Blaise; Bi, Guoqiang et al. (2021): Recommendations for Responsible Development and Application of Neurotechnologies. In: *Neuroethics* 14 (3), S. 365–386. DOI: 10.1007/s12152-021-09468-6.

Hertz, Nora (2023): Neurorights – Do we Need New Human Rights? A Reconsideration of the Right to Freedom of Thought. In: *Neuroethics* 16 (1). DOI: 10.1007/s12152-022-09511-0.

Hirsch-Kreinsen, Hartmut; Karacic, Anemari (2019): *Autonome Systeme und Arbeit*. Bielefeld, Germany: transcript Verlag. Online verfügbar unter <https://library.open.org/bitstream/handle/20.500.12657/25346/1/9783839443958.pdf>.

Ienca, Marcello; Ignatiadis, Karolina (2020): Artificial Intelligence in Clinical Neuroscience: Methodological and Ethical Challenges. In: *AJOB neuroscience* 11 (2), S. 77–87. DOI: 10.1080/21507740.2020.1740352.

Iliashenko, Oksana; Bikkulova, Zilia; Dubgorn, Alissa (2019): Opportunities and challenges of artificial intelligence in healthcare. In: *E3S Web Conf.* 110, S. 2028. DOI: 10.1051/e3sconf/201911002028.

Kehl, Christoph: Möglichkeiten und Grenzen ethischer Technikgestaltung. In: S. 151–170. Online verfügbar unter [https://www.researchgate.net/profile/Christoph-Kehl-2/publication/355208420\\_Moeglichkeiten\\_und\\_Grenzen\\_ethischer\\_Technikgestaltung/links/61ab79bdca2d401f27c4dff2/Moeglichkeiten-und-Grenzen-ethischer-Technikgestaltung.pdf](https://www.researchgate.net/profile/Christoph-Kehl-2/publication/355208420_Moeglichkeiten_und_Grenzen_ethischer_Technikgestaltung/links/61ab79bdca2d401f27c4dff2/Moeglichkeiten-und-Grenzen-ethischer-Technikgestaltung.pdf).

Larsson, Stefan; Heintz, Fredrik (2020): Transparency in artificial intelligence. In: *Internet Policy Review* 9 (2). DOI: 10.14763/2020.2.1469.

Lorenz, Jeanette Miriam (2022): Bessere KI-Algorithmen durch Quantencomputing? Online verfügbar unter <https://publica.fraunhofer.de/entities/publication/84e250b1-6691-4261-ac55-abf58c4406a7/details>.

Lubkowitz, Matthias (2022): Bessere KI-Algorithmen durch Quantencomputing? Online verfügbar unter <https://publica.fraunhofer.de/entities/publication/84e250b1-6691-4261-ac55-abf58c4406a7/details>.

McKinsey Global Institute (2018): NOTES FROM THE AI FRONTIER MODELING THE IMPACT OF AI ON THE WORLD ECONOMY. Online verfügbar unter <https://www.mckinsey.com/featured-insights/artificial-intelligence/notes-from-the-ai-frontier-modeling-the-impact-of-ai-on-the-world-economy>.

Müller, Oliver; Rotter, Stefan: *Neurotechnologie: Aktuelle Entwicklungen und ethische Fragen* 2016, S. 36–40. Online verfügbar unter <https://www.fiff.de/publikationen/fiff-kommunikation/fk-2016/fk-2016-2/fk-2016-2-content/fk-2-16-p36.pdf>.

Musk, Elon (2020): Neuralink Demo Event 2020. The Presentation, 29.08.2020.

Neurologie: Schalten die Gehirne Pädophiler einfach falsch? (2007). In: *WELT*, 30.11.2007. Online verfügbar unter <https://www.welt.de/wissenschaft/article1415943/Schalten-die-Gehirne-Paedophiler-ein-fach-falsch.html>, zuletzt geprüft am 04.05.2023.

Paek, Andrew Y.; Brantley, Justin A.; Evans, Barbara J.; Contreras-Vidal, Jose L. (2021): Concerns in the Blurred Divisions between Medical and Consumer Neurotechnology. In: *IEEE systems journal* 15 (2), S. 3069–3080. DOI: 10.1109/jsyst.2020.3032609.

Paszkiel, Szczepan; Rojek, Ryszard; Lei, Ningrong; Castro, Maria António (2021): A Pilot Study of Game Design in the Unity Environment as an Example of the Use of Neurogaming on the Basis of Brain–Computer Interface Technology to Improve Concentration. In: *NeuroSci* 2 (2), S. 109–119. DOI: 10.3390/neurosci2020007.

Patel, Vimla L.; Shortliffe, Edward H.; Stefanelli, Mario; Szolovits, Peter; Berthold, Michael R.; Bellazzi, Riccardo; Abu-Hanna, Ameen (2009): The coming of age of artificial intelligence in medicine. In: *Artificial intelligence in medicine* 46 (1), S. 5–17. DOI: 10.1016/j.artmed.2008.07.017.

Ponseti, Jorge; Granert, Oliver; Jansen, Olav; Wolff, Stephan; Beier, Klaus; Neutze, Janina et al. (2012): Assessment of pedophilia using hemodynamic brain response to sexual stimuli. In: *Archives of general psychiatry* 69 (2), S. 187–194. DOI: 10.1001/archgenpsychiatry.2011.130.

PwC (2023): How will automation impact jobs? Online verfügbar unter <https://www.pwc.co.uk/services/economics/insights/the-impact-of-automation-on-jobs.html>, zuletzt aktualisiert am 04.05.2023, zuletzt geprüft am 04.05.2023.

Rainey, Stephen; Erden, Yasemin J. (2020): Correcting the Brain? The Convergence of Neuroscience, Neurotechnology, Psychiatry, and Artificial Intelligence. In: *Science and engineering ethics* 26 (5), S. 2439–2454. DOI: 10.1007/s11948-020-00240-2.

Robinson, Jacob T.; Rommelfanger, Karen S.; Anikeeva, Polina O.; Etienne, Arnelle; French, Jennifer; Gelinas, Jennifer et al. (2022): Building a culture of responsible neurotech: Neuroethics as socio-technical challenges. In: *Neuron* 110 (13), S. 2057–2062. DOI: 10.1016/j.neuron.2022.05.005.

Santosh, KC (2023): AI, Ethical Issues and Explainability—Applied Biometrics. Hg. v. SpringerLink. Online verfügbar unter <https://link.springer.com/book/10.1007/978-981-19-3935-8>, zuletzt aktualisiert am 07.05.2023, zuletzt geprüft am 07.05.2023.

Savage, Neil (2019): How AI and neuroscience drive each other forwards. In: *Nature* 571 (7766), S15–S17. DOI: 10.1038/d41586-019-02212-4.

Stahl, Bernd Carsten; Schroeder, Doris; Rodrigues, Rowena (2022): Ethics of Artificial Intelligence. Case Studies and Options for Addressing Ethical Challenges. 1st ed. 2023. Cham: EU-funded Sherpa project.

Stuart Russell, Peter Norvig: Artificial Intelligence: A Modern Approach (3rd Edition). Online verfügbar unter <http://repo.darmajaya.ac.id/3800/1/Artificial%20Intelligence%20A%20Modern%20Approach%20%283rd%20Edition%29.pdf%20%28%20PDFDrive%20%29.pdf>.

Thompson, Neil C.; Greenewald, Kristjan; Lee, Keeheon; Manso, Gabriel F. (2021): Deep Learning's Diminishing Returns: The Cost of Improvement is Becoming Unsustainable. In: *IEEE Spectr.* 58 (10), S. 50–55. DOI: 10.1109/MSPEC.2021.9563954.

Topol, Eric J.; Verghese, Abraham (2019): Deep medicine. How artificial intelligence can make healthcare human again. First edition. New York, NY: Basic Books.

Unabhängige hochrangige Expertengruppe für Künstliche Intelligenz (2018): ETHIK-LEITLINIEN FÜR EINE VERTRAUENSWÜRDIGE KI.

Vukelic, Mathias (2023): Jedes Hirn tickt anders: Wie mit Neurotechnologie und KI berufliche Inklusion gelingen kann. Online verfügbar unter <https://blog.iao.fraunhofer.de/jedes-hirn-tickt-anders-wie-mit-neurotechnologie-und-ki-berufliche-inklusion-gelingen-kann/>, zuletzt aktualisiert am 13.03.2023, zuletzt geprüft am 04.05.2023