

Wolfgang Wahlster, Saarbrücken

Künstliche Intelligenz: Digitales Verstehen

Künstliche Intelligenz hat das Potential, die Forschungsmethodik in zahlreichen Fachdisziplinen zu revolutionieren. Ausgehend von dem automatischen Verstehen digitaler Daten kann mit maschinellem Lernen, wissensbasierten Assistenzsystemen und Laborrobotern eine bislang unerreichte Qualitätsstufe in der Forschung angestrebt werden. KI-Systeme analysieren die relevante Fachliteratur, führen selbstständig Experimente durch und analysieren deren Ergebnisse, erkennen neue Zusammenhänge in Massendaten und arbeiten formale Beweise für mathematische Behauptungen aus. Sie ersetzen aber nicht die Forschenden, sondern arbeiten unter deren Leitung und mit deren Hilfe gemeinsam an wissenschaftlichen Fragestellungen, die dem Menschen wichtig sind.

Nachdem die Mathematik als eine der ältesten Wissenschaftsdisziplinen sehr erfolgreich Methoden zur formalen Beschreibung und präzisen Analyse komplexer Strukturen in allen Bereichen der Wissenschaft geschaffen hat, entwickelte die Informatik auf diesen Grundlagen aufbauend auch formale Modelle für komplexe Prozesse, die auf digitalen Rechnern ausgeführt werden können. Beim Übergang von der Mathematik zur Informatik liegt der Schwerpunkt auf der mathematischen Analyse effizienter Algorithmen für die Daten- und Wissensverarbeitung – beispielsweise deren Zeit-, Rechen- und Speicherplatzbedarf – sowie den Grundlagen von Hard- und Software-Architekturen für Werkzeuge zur Analyse, Simulation und Ausführung komplexer Prozesse.

Nach der Mathematisierung der Wissenschaften seit mehreren Jahrhunderten, die große wissenschaftlichen Fortschritte ermöglichte, ist nun die Informatisierung seit ca. 50 Jahren in vollem Gange. Sie hat den Übergang von der präzisen Analyse von Strukturen zu Prozessen ermöglicht. Der Untertitel zum Leitthema *Digitalisierung der Wissenschaften* der 130. Versammlung der GDNÄ lautete daher zu Recht: *Von Strukturen zu Prozessen*.

Aus der Informatik heraus haben sich eine Vielzahl von digitalen Technologien wie Supercomputer, das Internet, mobile Computer und die Cloud entwickelt, welche die Grundlage für die erfolgreiche praktische Umsetzung der Digitalisierung der Wissenschaften bilden. Auf Mathematik, Informatik und digitalen Technologien aufbauend, hat sich seit 1956 mit der Künstlichen Intelligenz (KI, [1]) ein Teilgebiet der Informatik zu seiner heutigen Blüte entwickeln können, das erstmals ein digitales Verstehen wissenschaftlicher Struktur- und Prozessdaten ermöglicht (Abb. 1).

Damit ist die KI die Speerspitze der Digitalisierung. KI versucht als die Avantgarde der Informatik jeweils bestehende Grenzen der Digitalisierung zu durchbrechen. Nach dem Einzug in Fabriken, Büros, Verkehrsmittel und Wohnungen kommt

die KI jetzt auch zunehmend in die Forschungslabore. Immer mehr Wissenschaftler nutzen digitale Assistenzsysteme, um Routinearbeiten wie Recherchen und Datenanalysen, aber auch aufwendige Experimente automatisiert erledigen zu lassen. Das revolutioniert nicht nur das wissenschaftliche Arbeiten, sondern bietet auch die Chance, eine neue, bislang unerreichte Qualitätsstufe in der Forschung zu erklimmen.

KI-basierte Computersysteme stellen nicht mehr nur Daten digital bereit, verarbeiten und verteilen sie, sondern verstehen – zumindest in einigen Kontexten – auch deren Inhalt. So werden KI-Systeme zu digitalen Hilfswissenschaftlern. Durch rasante Fortschritte beim maschinellen Lernen können extrem große Datenmengen in kürzester Zeit gesichtet, klassifiziert und ausgewertet werden. Das bedeutet beispielsweise, dass in der Astrophysik Aufnahmen aus Teleskopen zehn Millionen Mal schneller als bisher ausgelesen werden oder dass das Zusam-

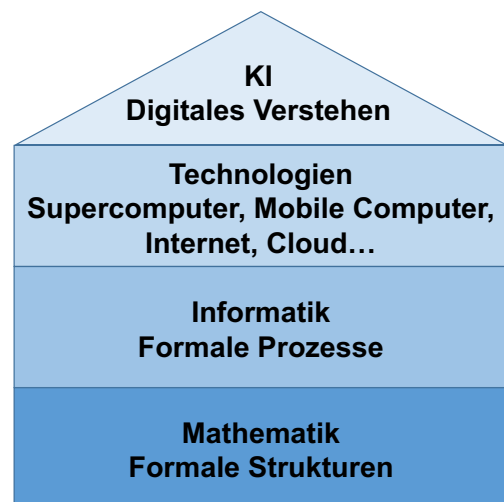


Abb. 1. Die Basis für das digitale Verstehen von Strukturen und Prozessen.

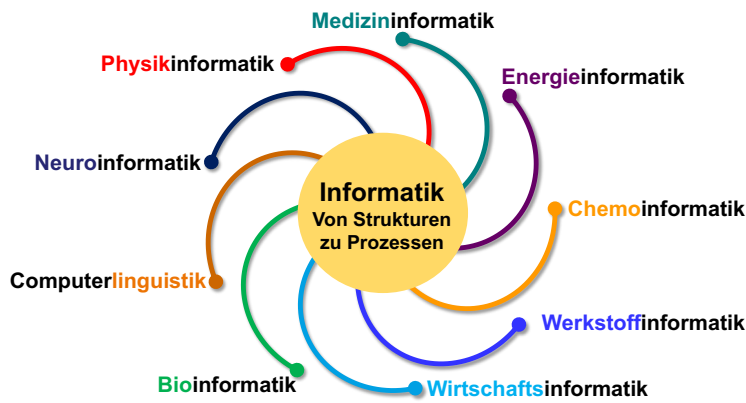


Abb. 2. Beispiele für die Informatisierung der Wissenschaften.

mensspiel von neuen Medikamenten mit Tausenden möglicher Nebenwirkungen vorausgesagt werden kann.

Zudem können Roboter-Teams verschiedener Spezialisierung in lebensfeindlichen Umgebungen wie im Weltraum oder in der Tiefsee Forschungsarbeiten autonom ohne Fernsteuerung erledigen und ihre Beobachtungen den Forschern melden. Schon heute arbeiten digitale Assistenten Hand in Hand mit den Forschern – zuverlässig und schneller als die Heerscharen von Hilfswissenschaftlern, die bisher die oft repetitive „Kärrnerarbeit“ in den Forschungslaboren leisten.

Informatische Methoden werden heute in fast allen Wissenschaften entwickelt, gelehrt und angewandt (Abb. 2). Es sind zahlreiche sog. Bindestrich-Informatiken entstanden, denen an denen besten Informatik-Standorten auch eigene Lehrstühle und Forschungsbereiche gewidmet sind. In den letzten Jahren war kaum noch ein Nobelpreis in Physik oder Chemie ohne den Einsatz digitaler Systeme zu gewinnen. Heute müsste die Behauptung von Immanuel Kant im Jahre 1781, dass „in jeder besonderen Naturlehre nur soviel eigentliche Wissenschaft angetroffen werden kann, als darin Mathematik anzutreffen ist“ durch „und Informatik angewandt wird“ ergänzt werden [2].

„Digitales Verstehen“ im Sinne des Untertitels dieses Beitrages eröffnet auf der Basis von KI aber viel weitergehende Möglichkeiten durch wissenschaftliche Assistenzsysteme. Der Titel wurde bewusst mehrdeutig gewählt, wobei zwei Lesarten für die weitere Digitalisierung der Wissenschaften relevant sind:

1. „Digitales Verstehen“ als Verstehen von digitalen Daten: Das Verb „verstehen“ wird bezogen auf das Objekt „Digitales“, also wissenschaftliche Daten, die digital erfasst und gespeichert sind.

2. „Digitales Verstehen“ als Verstehen mit Hilfe digitaler Computer: Das Adjektiv „digital“

modifiziert hier die Bedeutung des Nomens „Verstehen“, weil so der Unterschied zum menschlichen Verstehen betont wird.

Mit KI werden beide Lesarten kombiniert umgesetzt, nämlich das Verstehen von digitalen Daten mit der Hilfe von digitalen Computern.

Künstliche Intelligenz als Grundlage des digitalen Verstehens

Künstliche Intelligenz (KI) ist ein Teilgebiet der Informatik mit dem Ziel, intelligentes Verhalten und die zugrundeliegenden kognitiven Fähigkeiten auf digitalen Computern zu realisieren. Dabei erhebt man in der ingenieurwissenschaftlich ausgerichteten KI nicht den Anspruch, Funktionsmodelle zu implementieren, welche die menschliche Intelligenz exakt nachbilden. Man strebt stattdessen lediglich Leistungsmodelle an, die einige Problemlösungsfähigkeiten des Menschen approximieren und deren Ergebnisse reproduzieren. Die wissenschaftlichen Kerngebiete der KI sind neben der subsymbolischen Mustererkennung verschiedene, meist numerische Verfahren zum maschinellen Lernen, zur symbolischen Wissensrepräsentation und Wissensverarbeitung sowie zur Präsentation von Wissen für den menschlichen Benutzer (Abb. 3).

Für die Erforschung und Anwendung dieser Kerngebiete gibt es ein weites Spektrum von Werkzeugsystemen wie Wissenrepräsentationssprachen zur computergerechten Darstellung von Wissensgebieten, Ontologien zur Definition der Begrifflichkeit von Fachdomänen, Lern- und Inferenzbibliotheken als Baukasten, um spezielle maschinelle Lern- oder Schlussfolgerungsverfahren implementieren zu können und KI-Hardware wie TPU- oder GPU-Cluster, die extrem große Datenmengen hochparallel verarbeiten können (vgl. Abb. 3).

Neben KI-Systemen zum Sprach- und Bildverstehen sind beim Einsatz in den Wissenschaften auch autonome Systeme von großer Bedeutung, die z.B. im Weltall, der Tiefsee oder anderen für den Menschen lebensfeindliche Umgebungen

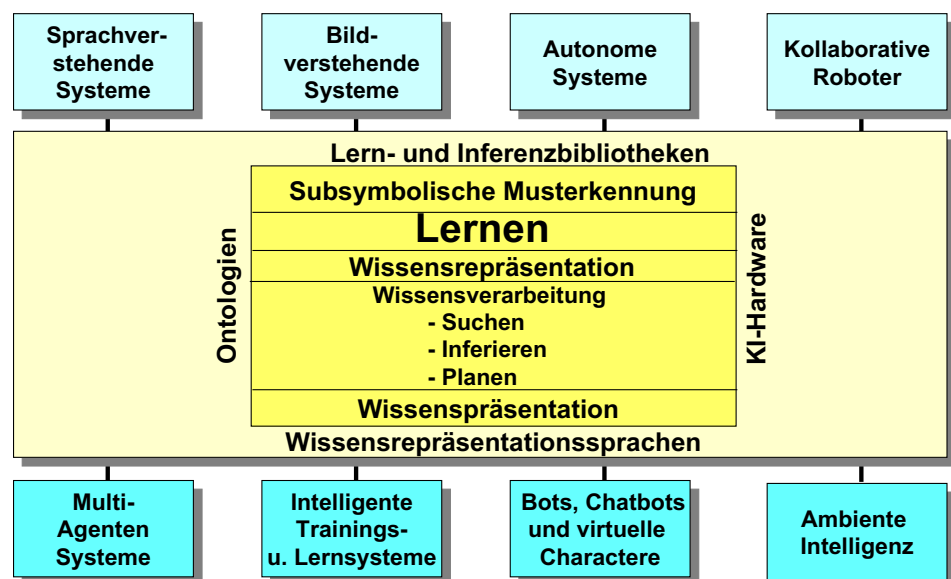


Abb. 3. Kerngebiete und Einsatzgebiete der KI.

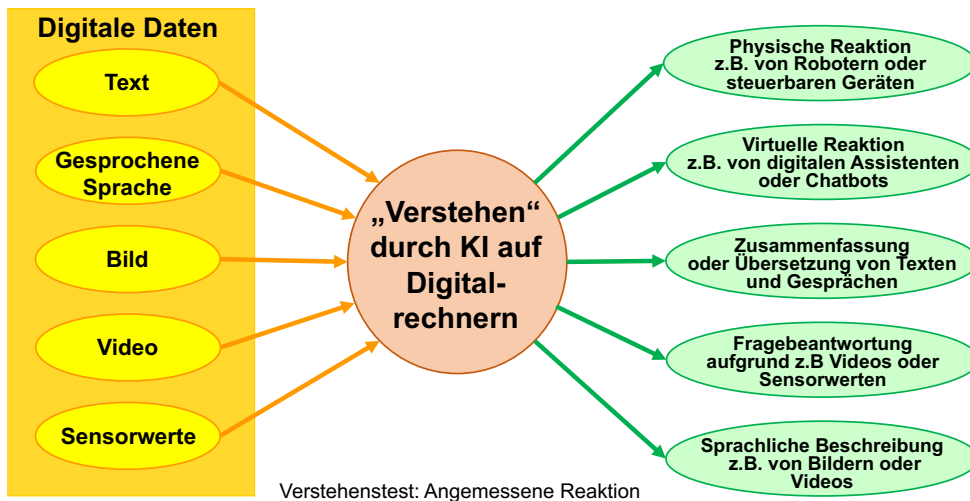


Abb. 4. Angemessene Reaktionen auf Eingabedaten als Verstehenstests für KI-Systeme.

Experimente ohne Fernsteuerung selbstständig ausführen. Aber auch kollaborative KI-Roboter, die Hand in Hand mit Wissenschaftlern arbeiten und ihnen einen Teil der physischen Arbeit abnehmen, sind heute schon im praktischen Einsatz.

Um zu testen oder nachzuweisen, dass ein KI-System digitale Daten verstanden hat, gibt es eine Vielzahl von möglichen Aufgaben, deren korrekte Lösung ein Verstehen der Eingabedaten voraussetzt (Abb. 4). Wie ein Lehrer, der überprüfen will, ob der Schüler einen Roman gelesen und dessen wesentlichen Inhalte verstanden hat, kann man auch von einem KI-System verlangen, eine Zusammenfassung zu generieren, die man dann auf Vollständigkeit und Korrektheit überprüft. Auch die Zusammenfassung oder gar korrekte Übersetzung von Gesprächsinhalten ist ein guter Verstehenstest [3].

Daneben können die sprachliche Bildbeschreibung [4] und die Beantwortung von Fragen durch Sprachdialogsysteme oder Chatbots als Evidenzen für ein digitales Verstehen gewertet werden. Besonders klar wird die Operationalisierung des Verstehensbegriffes in der KI bei der Interaktion mit Robotern und anderen autonomen Systemen [5]. Wenn ein kollaborativer Roboter aufgrund einer Spracheingabe: „Bitte reiche mir den Elektroböhrer an!“ dem Facharbeiter das richtige Werkzeug übergibt, gehen wir von einer positiven Verstehensleistung aus.

Grundlegend für den Verstehensbegriff, wie wir ihn in der KI verwenden, ist die Unterscheidung von Daten, Information, Wissen und Meta-Wissen (Abb. 5). Während Daten in der Informatik reine Zeichenfolgen sind, die durch eine Interpretation im situativen Kontext für einen Empfänger der Daten in einer Nachricht erst zu einer spezifischen Information werden,

führt erst Wissen zum Nutzen von verknüpfter Information für Problemlösungs-, Entscheidungs- oder Schlussfolgerungsprozesse zur praktischen Anwendung. Meta-Wissen als Wissen über Wissen erweist sich als äußerst wertvoll für KI-Systeme, wenn es beispielsweise um die Nutzung inhärent unsicherer Informationsquellen geht. Für alle genannten vier Ebenen haben wir in der Informatik prozedurale Methoden entwickelt, um diese für KI-Systeme nutzbar zu machen. Die binäre Codierung der Zahl 50 kann dann im Auto als Geschwindigkeitsangabe

interpretiert werden. Aufgrund des sicheren Wissens aus dem Fahrschulunterricht, dass dies die Höchstgeschwindigkeit in geschlossenen Ortschaften ist, kann man den Autofahrer durch ein KI-basiertes Assistenzsystem warnen lassen, wenn er diesen Richtwert bei Einfahrt in ein Stadtgebiet überschreitet. Digitales Verstehen von Daten kann also zu neuer Information, zu Wissen oder sogar Meta-Wissen für ein KI-System führen.

Das Beispiel der Verkehrsregel zeigt auch sehr klar, dass maschineller Wissenserwerb nicht nur auf empirischem Lernen über Massendaten beruhen kann, sondern auch über die Informationsextraktion aus der Fachliteratur oder interaktives Lernen im Dialog mit einem menschlichen Wissensträger als Lehrer erfolgen muss (Abb. 6). Nur durch Beobachtung der Geschwindigkeiten von Verkehrsteilnehmern in geschlossenen Ortschaften kann nämlich auch bei einer noch so großen Datenbasis kaum der exakte Wert 50 km/h induziert werden. Denn kaum ein Fahrer hält exakt diesen Wert ein, so dass eher ein Durchschnittswert wie 54,5 km/h gelernt würde, was aber nicht die präzise juristische Norm ist, die man nur durch Studium der Regelwerke oder durch Interaktion mit einem Fahrlehrer als Wissen erwerben kann.

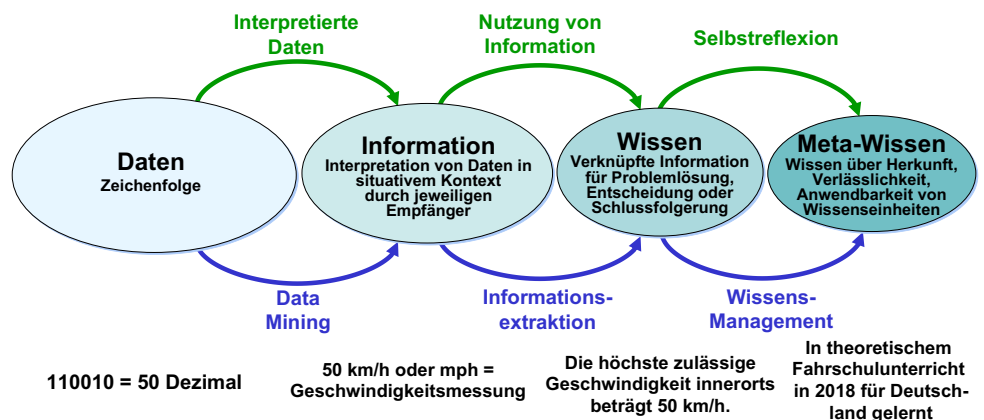
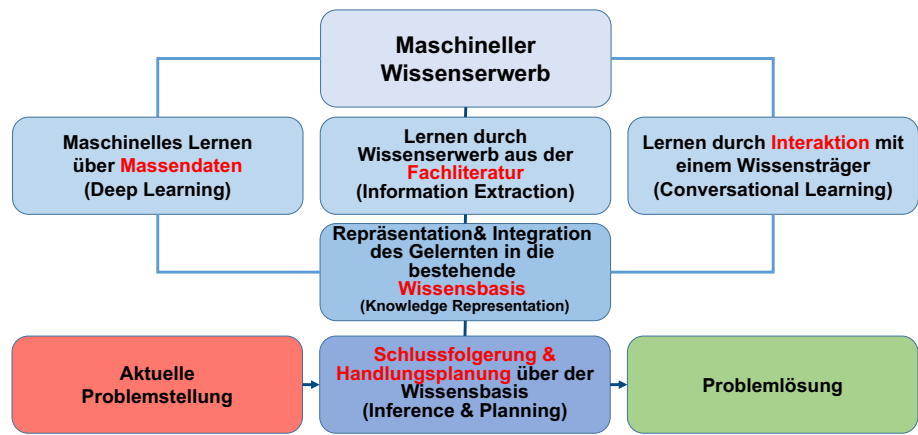


Abb. 5. Die Kaskade digitaler Verstehensprozesse von Daten bis zum Meta-Wissen.

Abb. 6. Drei verschiedene Zugänge zum Wissenserwerb für KI-Systeme.



Wie Abbildung 6 zeigt, gehört auch die Repräsentation und die Integration des Gelernten in die bereits bestehende Wissensbasis zu den Kernaufgaben eines erfolgreichen Wissenserwerbs für ein KI-System. Erst dann kann das erworbene Wissen durch das KI-System dazu genutzt werden, eine aktuelle Problemstellung zu lösen und sein Verstehen digitaler Daten zu optimieren. Bei der aktuell weitverbreiteten ausschließlichen Fokussierung auf maschinelles Lernen über Massendaten werden die beiden anderen Arten des Wissenserwerbs häufig durch KI-Neueinsteiger vernachlässigt oder sogar übersehen. Das ist aber für das digitale Verstehen in den Wissenschaften völlig inadäquat, da dort gerade die Kommunikation zwischen Experten von größter Bedeutung für den wissenschaftlichen Fortschritt ist.

Physische und kognitive Assistenzsysteme für die Wissenschaft

Die zweite Welle der Digitalisierung wird geprägt durch die neuen Möglichkeiten des digitalen Verstehens. Die erste Welle der Digitalisierungen führte zur umfassenden Ablösung analoger Methoden und zum nahezu vollständigen digitalen Erfassen, Speichern, Übertragen und Speichern von Daten. Kein Forschungslabor kommt heute mehr ohne digitale Speicher, Rechner und Internetverbindungen aus. Die Internet- und Cloud-Technologien sind in der Wissenschaft nicht mehr wegzudenken. Alle Daten sind maschinenlesbar. Aber die zweite Welle der Digitalisierung, die auf KI und deren Beiträge zum Maschinellen Lernen beruht, führt in die neue Ära der maschinenverstehbaren Daten. Dabei werden die Daten inhaltlich interpretiert, klassifiziert, angereichert mit Meta-Daten und veredelt, um dann neue Schlussfolgerungen daraus ziehen zu können (Abb. 7).

Während bisher überwiegend Menschen die Interpretation digitaler Daten meist an Bildschirmarbeitsplätzen durch-

geführt haben, wird es jetzt vermehrt möglich, die Interpretation der digitalen Daten und ausgewählte Entscheidungen auf geringe Kritikalitätsstufen durch KI-Systeme vornehmen zu lassen (Abb. 8).

Die bereits abgeschlossene erste Welle der Digitalisierung der Wissenschaften hatte bereits wichtige Auswirkungen auf den Alltag und die Produktivität von Wissenschaftlern:

- Weltweite Literatur- und Faktensuche im World Wide Web mit Übersetzungsunterstützung bei fremdsprachigen Literaturquellen
- Digitale Publikation und beschleunigte weltweite wissenschaftliche Kommunikation
- Digitale Visualisierung von Massendaten, Struktur- und Prozessmodellen
- Algorithmische Beschreibung und Analyse komplexer Prozesse
- Wissenschaftliche Softwarepakete zur freien Nutzung und Weiterentwicklung
- Verteilte Datensammlung auch mit der Hilfe interessierter Laien (Crowd Sourcing, Citizen Science)
- Überprüfung von Ergebnissen und Leistungsvergleich von Verfahren durch Zugriff auf standardisierte digitale Datensätze (Benchmarking).

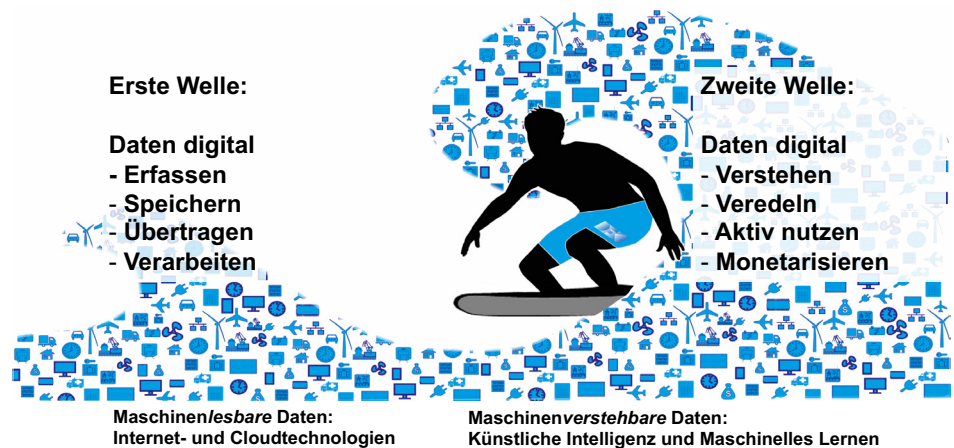


Abb. 7. Die zwei Wellen der Digitalisierung.

Die zweite Welle der Digitalisierung, die zunehmend auf Künstlicher Intelligenz beruht, bietet als Fähigkeitsverstärker für Forschende aber viel weitgehendere Möglichkeiten für die Wissenschaft.

Das gilt für die Unterstützung von körperlicher und geistiger Arbeit in Forschungsvorhaben durch KI-basierte Assistenzsysteme. Physische Assistenz können z.B. robotische KI-Systeme anbieten, während Softwarebasierte KI-Systeme eine kognitive Assistenz ermöglichen.

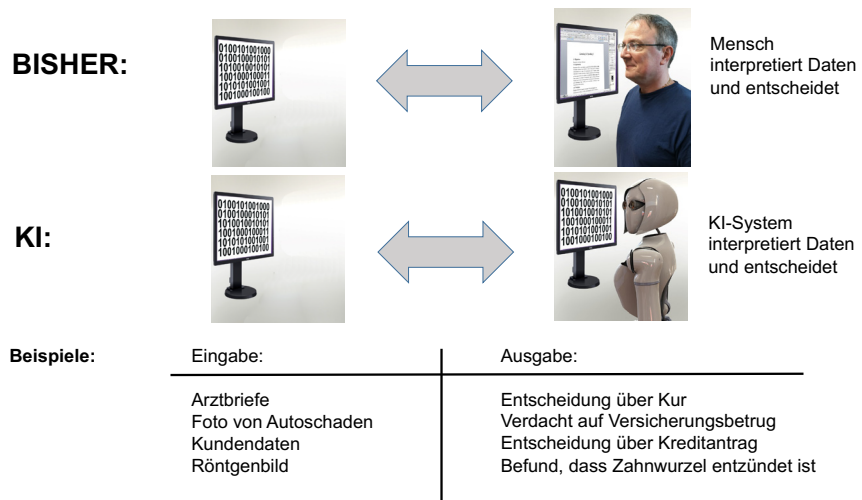


Abb. 8. Menschliches versus Maschinelles Auswerten digitaler Daten.

Wichtige Beispiele für physische Assistenzsysteme sind:

- Die robotische Hochdurchsatzforschung (z. B. Wirkstoff-, Werkstoff-Screening)
- KI-basierte Steuerung komplexer wissenschaftlicher Instrumente (z. B. Teleskope oder Mikroskope)
- Robotische Beobachtungen und Probenentnahme in lebensfeindlichen Umgebungen (z. B. Tiefsee, Weltraum, kontaminierte/einsturzgefährdete Bauten)
- Drohnen als mobile Sensorplattformen (z. B. in Archäologie)
- Robonauten als Assistenzsysteme für Astronauten
- Exoskelette zur Fernerkundung und Steuerung von entfernten Aktuatoren.

Bei der kombinatorischen Wirkstoffforschung wird der quantitative Fortschritt durch den Einsatz von Laborrobotern bei der Suche nach neuen Medikamenten besonders deutlich. Pharmaunternehmen kombinieren heute mit Robotern 1,6 Millionen chemische Verbindungen pro Tag im 24-Stunden-Betrieb und erhalten so jeden Tag rund 150 000 konsistente Datenpunkte für die Entwicklung neuer Medikamente. Vor dem Einsatz solcher robotischer Hochdurchsatzsysteme konnte man dagegen durch die physische Arbeit von Laborpersonal täglich lediglich 30 bis 40 relevante Datenpunkte finden. Da durch physische Assistenzsysteme extrem große Datensätze (sog. Big Data) generiert werden, die manuell kaum noch auswertbar sind, werden zusätzlich kognitive Assistenzsysteme notwendig, die bei der Datenanalyse helfen.

Wichtige Beispiele für kognitive Assistenten als „digitale Hilfswissenschaftler“ sind:

- Maschinelles Lernen zur Datenanalyse, Klassifikation, Prognose, Hypothesen- und Modellbildung über Big Data
- Automatische Informationsextraktion aus Bildern, Videos, Texten, akustischen Daten und Sensordaten
- Automatische Deduktion als mathematischer Beweisassistent und zur Verifikation von Modellen
- KI-basierter Überblick zum Forschungsstand und Identifikation von Forschungstrends
- Virtualisierung und Simulation von Experimenten in digitaler Realität und erweiterter Realität als Forschungsassistenz.

Neben der Datenanalyse können auch andere wichtige Teilleistungen der wissenschaftlichen Arbeit, wie die Ausarbeitung eines formalen Beweises in der Mathematik von KI-Systemen unterstützt werden. So ist es mit Hilfe des französischen Beweissystems Coq gelungen, den ersten voll formalen Beweis für den Vierfarbensatz zu finden [6]. Wie bei der Hochdurchsatzanalyse entstehen auch beim automatischen Beweisen in vielen Fällen Datensätze, die von Menschen kaum noch überprüfbar sind: So wurde beispielsweise für einen Spezialfall der Diskrepanz-Vermutung von Erdős mithilfe eines KI-Assistenzsystems ein Beweis von 13 Gigabyte Länge gefunden, der ca. 1 Milliarde einfache Schlussfolgerungen umfasste. Nur durch ein algorithmisches Metaprüfsystem konnte dieser extrem lange Beweis für menschliche Mathematiker verständlich gemacht und als korrekt anerkannt werden.

Derzeit wird aber das maschinelle Lernen als das wichtigste Teilgebiet der Künstlichen Intelligenz angesehen, wenn es darum geht, KI-Systeme in den wissenschaftlichen Erkenntnisprozess einzubetten, um grundlegende neue Prinzipien aus Daten abzuleiten

Maschinelles Lernen für die Wissenschaft

In den Anfängen der KI-Forschung bis in die 80er Jahre wurden zahlreiche Systeme direkt durch Programmierer auf der Basis heuristischer Algorithmen in symbolverarbeitenden Programmiersprachen wie LISP umgesetzt. Wie bei jeder Software war der Entwicklungsaufwand hoch und die Adaption der Lösung bei veränderten Randbedingungen aufwendig. Bei dieser Vorgehensweise war es auch schwer, die Ergebnisse der KI-Systeme automatisiert erklären zu lassen, da oft nur der Programmierer selbst seinen Programmcode verstehen und erläutern konnte. Ein digitales Verstehen des Programmcodes war für komplexe Software nicht möglich. Da aber die Erklärungsfähigkeit von digitalen Assistenten besonders in der wissenschaftlichen Anwendung fundamental für deren Akzeptanz ist, hat man in der nächsten Phase der KI-Entwicklung wissenbasierte Systeme entwickelt, in denen in einer Wissensbasis Fakten, Regeln und Modelle deklarativ und in einer formalen Wissensrepräsentationssprache mit präziser Semantik gespeichert

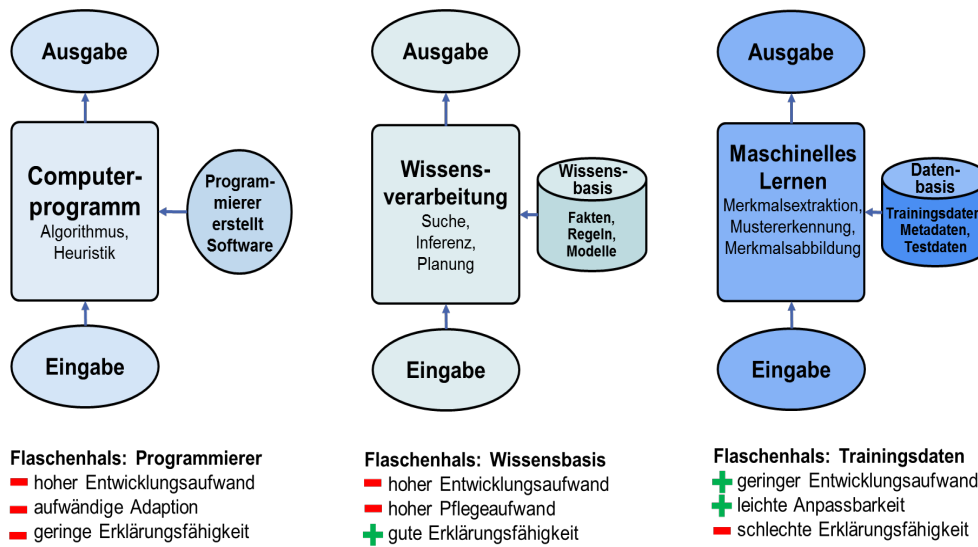


Abb. 9. Von der Programmierung zum Maschinellen Lernen.

werden. Das KI-System interpretiert und verknüpft Elemente der Wissensbasis automatisch so, dass eine Lösung gefunden werden kann (Abb. 9).

Wissensbasierte Systeme bieten eine sehr gute Basis für eine automatisierte Erklärung der Verarbeitungsergebnisse, weil sie ja expliziten Zugriff auf das von ihnen verwendete Wissen haben. Die Wissensbasen sind aber nur mit großem Aufwand manuell von Fachexperten für eine Anwendungsdomäne zu erstellen und zu pflegen.

Wenn man genügend Daten hat, kann man sich die Programmierung oder den Aufbau einer Wissensbasis sparen, da durch maschinelles Lernen in vielen Fällen automatisiert eine Abbildung der Eingabe auf die intendierte Ausgabe gefunden werden kann (Abb. 9). Der Entwicklungsaufwand wird damit erheblich verringert, und bei einer neuen Datenlage kann durch weiteres Lernen eine Anpassung des Verfahrens erreicht werden. Derzeit ist allerdings die Erklärungsfähigkeit der meisten Lernverfahren noch gering. Hierzu wird derzeit intensiv an Verfahren gearbeitet, welche die Ergebnisse von Lernalgorithmen für einen Menschen verständlich erklären können, um dessen Vertrauen in die Ergebnisse des Lernsystems sicherzustellen.

Statistisches Lernen wird heute vorwiegend über ein großes Spektrum von Lernarchitekturen mit neuronalen Netzen realisiert. Diese spielten lange eine Nebenrolle in der KI, obwohl selbst auf der ersten KI-Konferenz im Jahr 1956 schon einfachste Anwendungen neuronaler Netze diskutiert wurden, z. B. um geschriebene oder gesprochene Ziffern von 0 bis 9 zu erkennen.

Ein neuronales Netz ist ein System miteinander verbundener künstlicher Neuronen, die Nachrichten untereinander austauschen. Die Verbindungen haben eine numerische Gewichtung, die während des Trainingsprozesses angepasst wird, so dass ein korrekt trainiertes Netzwerk bei einem zu erkennenden Muster richtig reagiert. Das Netzwerk besteht aus mehreren Schichten mit Merkmal-erkennenden Neuronen. Jede Schicht verfügt über eine Vielzahl an Neuronen, die auf verschiedene Kombinationen von Eingaben von den vorhergehenden Schichten reagieren.

Die Schichten sind so aufgebaut, dass die erste Schicht verschiedene primitive Muster in der Eingabe entdeckt, die zweite Schicht Muster von Mustern, die dritte Schicht entdeckt Muster jener Muster und so weiter bis dann in der Ausgabeschicht das spezifische Ergebnis der Klassifikation erscheint. In Abbildung 10 wird dieser Schichtenaufbau am Beispiel der Klassifikation eines Eingabebildes und die Generierung einer passenden sprachlichen Bildbeschreibung als Ausgabe „Unfall auf Landstraße“ veranschaulicht. Beim Deep Learning werden aber die in einer Schicht zu analysierenden Eingabemerkmale nicht wie bei der klassischen Mustererkennung vorgegeben, sondern auch vom System selbst in den Lernzyklen bestimmt.

Das sogenannte überwachte Training erfolgt mit einem Datensatz, der eine große Zahl von repräsentativen Eingabemustern enthält, die mit der jeweils erwarteten Ausgabe-Antwort erweitert werden. Das Training mit annotierten Daten wird dazu genutzt, die passende Gewichtung der Neuronen für die Zwischen- und Endmerkmale iterativ zu bestimmen.

Neuronale Netze sind zwar biologisch von Gehirnen – insbesondere in den Grundbausteinen und der Terminologie – inspiriert, sind aber eher Graph-basierte Berechnungsmodelle als eine Gehirnsimulation, da das Gehirn erheblich komplexer strukturiert ist. Das Nervensystem ist beim Menschen außerdem mit dem Hormonsystem gekoppelt und damit dessen biochemischen Botenstoffen ausgesetzt, die Informationsverarbeitungsprozesse im Nervensystem auch durch Emotionen über das Hormonsystem beeinflussen können. Dies ist bei digitalen Computer selbstverständlich nicht möglich, da diese über kein Analogon zum Hormonsystem verfügen.

Ein Durchbruch wurde in den letzten Jahren mit dem Deep Learning [7] erreicht, das auf einer Weiterentwicklung mehrschichtiger neuronaler Netze beruht. Umso mehr verdeckte Schichten ein Deep Learning-Netz hat, umso mehr kann es abstrahieren und damit auch komplexere Zusammenhänge und subtile Merkmale in den Daten erkennen. Allerdings wird mit jeder Schicht auch die erforderliche Rechenleistung dramatisch erhöht. Es wird bereits an Netzen mit über 1000 Schichten gearbeitet, um Forschungsprobleme in der Medizin, der Physik, der

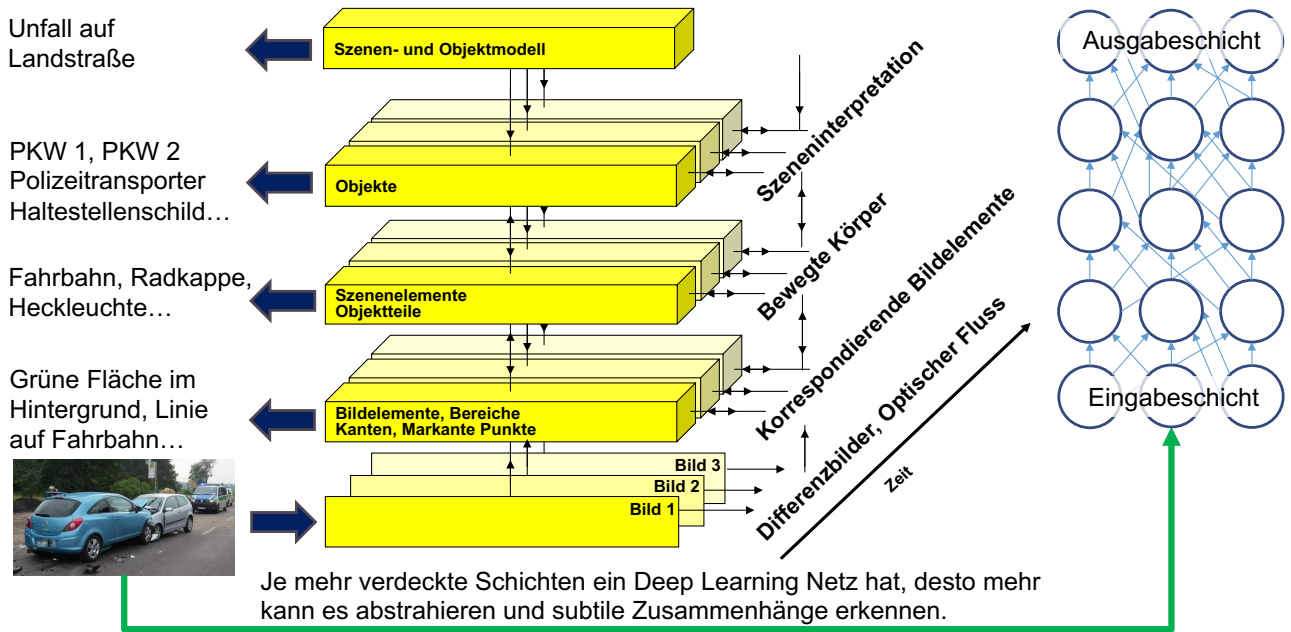


Abb. 10. Deep Learning mit mehrlagigen Netzen für kaskadierte Abstraktionsstufen.

Biologie und Chemie zu lösen. Besonders die Konvolutionsnetze (CNNs) haben aufgrund ihrer Effizienz viele neue Durchbrüche der KI ermöglicht. Deren Anfangsschichten analysieren jeweils nur kleine Ausschnitte der Eingabedaten als rezeptive Felder, die sich aber überlappen, so dass Kontextphänomene, wie sie u.a. für das Sprach- und Bildverstehen charakteristisch sind, trotz der erforderlichen Informationsverdichtung berücksichtigt werden.

Der Erfolg von Deep Learning hängt auch mit der Möglichkeit zur massiven Parallelisierung und damit auch der Beschleunigung des Trainings zusammen. Computergraphikkarten - sogenannte GPUs (Graphics Processing Units) eignen sich zur Parallelisierung des Trainings neuronaler Netze. GPUs haben massiv-parallele Architekturen, die eine fast 100fache Beschleunigung der Berechnungen beim Maschinellen Lernen bewirken. Wenn die auf einem großen GPU-Cluster trainierten neuronalen Netze in einer weiteren Verarbeitungsphase optimiert und kompiliert werden, können auf die Weise lauffzeitoptimierte Netze auch auf den Graphikkarten mobiler und eingebetteter Computer für Laboranwendungen zu geringen Kosten ausgeführt werden (Abb. 11). Damit können auch kleine Labore und Doktoranden z.B. mit Cloud-basierten Diensten den ersten Schritt des maschinellen Lernens über ihren Daten extern durchführen und das Resultat als optimiertes Netz dann für ihre Arbeiten auf kleineren internen Systemen zur lokalen Anwendung bringen.

Selbst in der Quantenphysik können mit Deep Learning in Echtzeit ganz neue Effekte analysiert werden, die sonst kaum zugänglich sind. Laserphysikern

der Universität Hamburg ist es gelungen, mit Hilfe von maschinellem Lernen aus experimentellen Daten Quantenphasenübergänge zu erkennen (Abb. 12), also die Punkte, an denen sich Eigenschaften von Stoffen ändern [8].

Bei ihren Experimenten fangen die Physiker ultrakalte Atome in einem Gitter aus Laserlicht und simulieren damit die Physik der Elektronen in einem Festkörper. Im Experiment gewonnene Bilder der Impulsverteilung werden nach erfolgreichem überwachten Lernen mit hoher Konfidenz der jeweils richtigen Phase zugeordnet, um damit die Phasenübergänge zu lokalisieren.

Die Hybride Kombination von statistischen und wissenschaftsbasierten KI-Methoden

In den letzten Jahren hat sich aber gezeigt, dass eine rein auf maschinellem Lernen basierende Künstliche Intelligenz schnell an Grenzen stößt. Das von der Menschheit über Jahrtausende

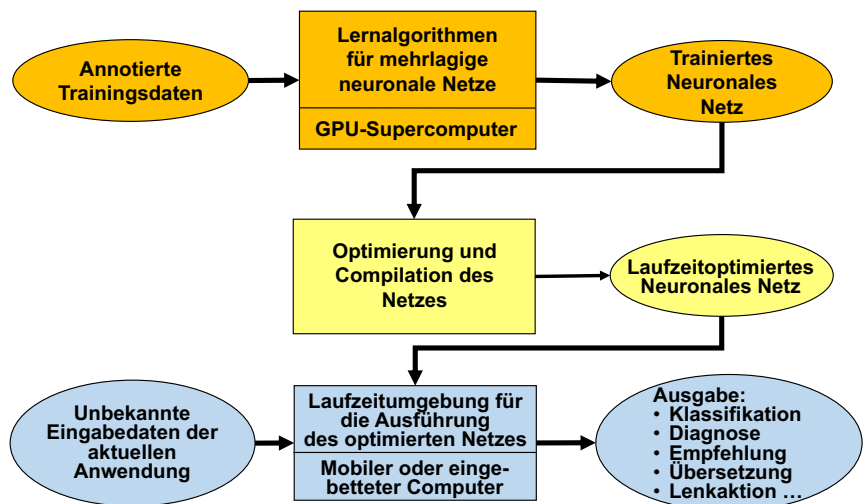


Abb. 11. Trainingsphase und Anwendungsphase neuronaler Netze.

de akkumulierte Wissen und das darauf basierende intelligente Verhalten lassen sich kaum durch statistisches Lernen über Beobachtungsdaten maschinell ersetzen. Es macht auch keinen Sinn, den Computer durch eigene physikalische Experimente Daten gewinnen zu lassen, um z.B. längst bekannte Zusammenhänge wie die Maxwell'schen Gleichungen mithilfe von Lernverfahren neu herzuleiten, anstatt den KI-Systemen bewährtes Wissen ohne Trainingsaufwand in deren Wissensbasis bereitzustellen.

Derzeit gibt es trotz der großen Erfolge beim reinen maschinellen Lernen auch noch zahlreiche offene Probleme:

- Wie kann auch bei sehr wenigen Trainingsdaten erfolgreich gelernt werden?
- Wie kann der Aufwand für die Annotation der Trainingsdaten reduziert werden?
- Wie kann durch Löschlernen auf kontaminierte Trainingsdaten reagiert werden?
- Wie kann eine Überanpassung bei zu langer Trainingsdauer verhindert werden?
- Wie können Scheinkorrelationen entdeckt und beseitigt werden?
- Wie können die schwachen Erklärungsleistungen umfassend verbessert werden?
- Wie kann die Robustheit der gelernten Netze gegenüber geringfügigen Transformationen der Trainingsdaten garantiert werden?
- Wie kann die Suche nach einer passenden Lernarchitektur theoretisch fundiert werden?
- Wie können vorhandene Wissensgraphen zur Validierung von Lernergebnissen eingesetzt werden?
- Wie kann verteiltes Lernen in vernetzten eingebetteten Systemen realisiert werden?
- Wie können vorhandene Modelle und Domänenwissen mit Lernsystemen verbunden werden?

Aktuell setzt man daher in der jüngsten Phase der Künstlichen Intelligenz vermehrt auf hybride Verfahren. Dabei werden Wissensbasen, die man z.B. aus der Informationsextraktion von Textdokumenten erstellt, kombiniert mit dem Lernen über Trainingsdaten – also eine Kombination von Methoden von KI-Systemen des maschinellen Lernens mit wissenbasierten Verfahren.

Dies ermöglicht eine wechselseitige Unterstützung von wissensfreien KI-Methoden, also statistischen Lernverfahren wie tiefen neuronalen Netze, und wissenbasierten KI-Methoden,

also Such-, Planungs- und Inferenzverfahren wie die terminologischen Logiken.

Statistische Lernverfahren können genutzt werden,

- um Suchprozesse in wissensbasierten Verfahren zu steuern und damit effizienter zu machen und
- um Grundoperatoren für wissensbasierte Inferenz- und Planungsverfahren zu lernen.

Wissenbasierte KI-Verfahren können genutzt werden,

- um die Ergebnisse von maschinellen Lernverfahren zu kombinieren, zu vervollständigen oder zu korrigieren und
- um die Ergebnisse aus maschinellen Lernverfahren plausibel zu machen oder zu erklären.

Bei Anwendungen mit hoher Systemkritikalität (z.B. medizinische Diagnose) spielen aus ethischer und juristischer Sicht besonders die Zuverlässigkeit der Resultate, die für den Anwender nachvollziehbare Erklärung und Rechtfertigung der Ergebnisse der KI-Systeme und die Robustheit gegenüber Systemstörungen eine besondere Rolle. Ethische Randbedingungen müssen schon beim Entwurf von KI-basierten Assistenzsystemen für Wissenschaftler unbedingt berücksichtigt werden (ethics-by-design).

Ausblick: Menschzentrierte KI versus maschinelle Superintelligenz

Zwar kritisieren seriöse KI-Forscher immer wieder Singularitäts-Propheten und Transhumanismus-Anhänger, die eine dem Menschen überlegene Superintelligenz als Vision propagieren. Sie schüren damit die Angst vor der Unterwerfung der Menschheit durch „superintelligente“ Roboter, wie sie aus trivialen Science Fiction-Romanen bekannt sind. Aber leider finden solche Scharlatane in den Massenmedien immer wieder Gehör. Dagegen haben die in dieser Arbeit präsentierten Ansätze das klare Ziel, eine die menschlichen Intelligenz unterstützende, ergänzende oder auch komplementäre maschinelle Intelligenz zu entwickeln, so dass menschlicher und maschineller Intellekt zusammen Probleme lösen, welche die Menschheit beschäftigen. Es sollen dabei durchaus auch bekannte Defizite menschlicher Intelligenz durch künstliche Intelligenz kompensiert werden. Das bringt uns im gesellschaftlichen Diskurs besser voran als die Diskussionen um eine „Superintelligenz“. Wir arbeiten

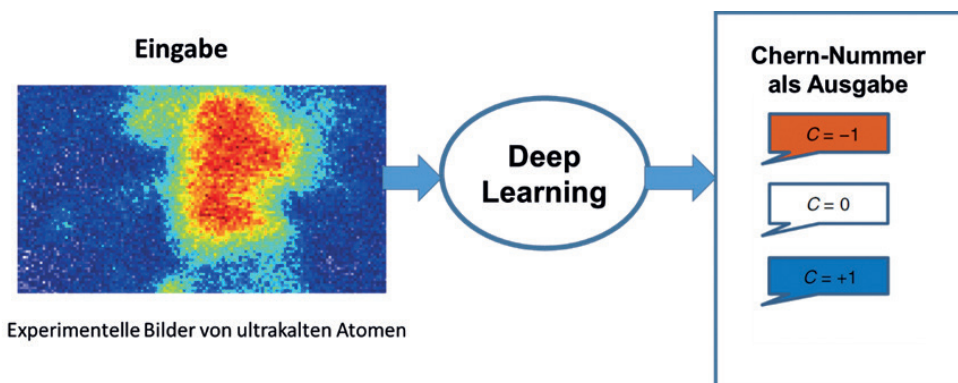


Abb. 12. Identifikation von Quantenphasenübergängen durch maschinelles Lernen. [Photo AG Sengstock/ UHH]

als KI-Wissenschaftler also auf keinen Fall an einem Homunkulus, sondern an Assistenzfunktionen für den Menschen, der stets im Mittelpunkt unserer Arbeit steht. Wie bei jeder Hochtechnologie besteht die Hauptgefahr darin, dass KI-Systeme von Kriminellen, Terroristen oder totalitären Regimen genutzt werden, um ihre menschenfeindlichen Ziele zu verfolgen.

Forschende sollen und können durch unsere menschenzentrierten KI-Systeme keinesfalls ersetzt werden, sondern sie sollen durch „Digitales Verstehen“ so unterstützt werden, dass der wissenschaftliche Fortschritt weiter beschleunigt und die Qualität der Forschung erhöht werden kann.

Literatur

[1] S. Russell, P. Norvig: Künstliche Intelligenz. 3. Auflage. Pearson. München 2012. – [2] I. Kant: Metaphysische Anfangsgründe der Naturwissenschaft, neu herausgegeben von Alois Höfler. Vorrede. Verlag von C. E. M. Pfeffer. Leipzig 1900. – [3] W. Wahlster: Mobile Speech-to-Speech Translation of Spontaneous Dialogs: An Overview of the Final Verbmobil System. In: Wahlster (Hrsg.): Verbmobil: Foundations of Speech-to-Speech Translation. Springer. Berlin, Heidelberg, New York 2000. – [4] W. Wahlster: Mit den Dingen sprechen: Autos, Roboter und Weinflaschen als Dialogpartner? In: T. Lengauer (Hrsg.): Computermodelle in der Wissenschaft – zwischen Analyse, Vorhersage und Suggestion. Nova Acta Leopoldina NF **110** Nr. 377, 119 (2011). – [5] W. Wahlster: Künstliche Intelligenz als Grundlage autonomer Systeme. Informatik Spektrum **40** (5), 409 (2017). – [6] G. Gonthier: Formal Proof – The Four-Color Theorem. Notices of the American Mathematical Society. **55** (11), 1382 (2008). – [7] M. Deru, A. Ndiaye: Deep Learning mit TensorFlow, Keras und Tensor Flow.js. Rheinwerk. Bonn 2019. – [8] B. S. Rem et al.: Identifying Quantum Phase Transitions using Artificial Neural Networks on Experimental Data. Nature Physics **15**, 917 (2019).



Prof. Dr. rer. nat. Dr. h. c. mult. **Wolfgang Wahlster** (Jahrgang 1953) studierte Informatik und Theoretische Linguistik an der Universität Hamburg. Im Anschluss an die Promotion in Informatik folgte er 1982 einem Ruf auf eine Professur für Informatik an der Universität des Saarlandes, der er bis zur Emeritierung (von 1985 ab als Professor für Informatik mit Lehrstuhl für Künstliche Intelligenz) verbunden blieb. Neben seiner Hochschultätigkeit

war Wahlster ab 1988 der Gründungsdirektor und ab 1997 bis Januar 2019 Vorsitzender der Geschäftsführung und technisch-wissenschaftlicher Leiter des Deutschen Forschungszentrums für Künstliche Intelligenz GmbH (DFKI).

Einer seiner Forschungsschwerpunkte sind Kognitive Assistenzsysteme. Er gehört zu den Mitschöpfern des Zukunftsprojekts Industrie 4.0, das 2011 erstmals veröffentlicht wurde. Wahlster ist auch nach seiner Emeritierung weiterhin Chefberater des DFKI und bleibt vielfach engagiert, u.a. als Mitglied des Lenkungskreises der BMBF-Plattform „Lernende Systeme“, der Forschungsplattform Industrie 4.0, der Datenethikkommission der Bundesregierung und als Vorsitzender des Steuerungskreises für die KI-Normungsroadmap. Für seine Arbeiten wurde er u.a. mit dem Deutschen Zukunftspreis (2001) und der Verleihung der Ehrendoktorwürde durch die Universität Linköping (1998), die Technischen Universität Darmstadt (2001) und die Universität Maastricht (2015) sowie das Große Verdienstkreuz (2019) ausgezeichnet. Wahlster ist Mitglied der Mainzer Akademie der Wissenschaften und Literatur (Mainz), der Königlich Schwedischen Akademie der Wissenschaften, der Leopoldina, der Deutschen Akademie der Technikwissenschaften acatech und der Berlin Brandenburgischen Akademie der Wissenschaften. Von 2017 bis 2018 war er Präsident der GDNÄ und organisierte die 130. Versammlung in Saarbrücken.

Deutsches Forschungszentrum für Künstliche Intelligenz GmbH, Alt-Moabit 91c, 10559 Berlin