

Carsten Lanquillon
Sigurd Schacht

Knowledge Science – Grundlagen

Methoden der Künstlichen Intelligenz für
die Wissensextraktion aus Texten

 Springer Vieweg



Knowledge Science – Grundlagen

Carsten Lanquillon · Sigurd Schacht

Knowledge Science – Grundlagen

Methoden der Künstlichen Intelligenz für
die Wissensextraktion aus Texten

Carsten Lanquillon
Hochschule Heilbronn
Heilbronn, Baden-Württemberg, Deutschland

Sigurd Schacht
Hochschule Ansbach
Ansbach, Baden-Württemberg, Deutschland

ISBN 978-3-658-41688-1 ISBN 978-3-658-41689-8 (eBook)
<https://doi.org/10.1007/978-3-658-41689-8>

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

© Der/die Herausgeber bzw. der/die Autor(en), exklusiv lizenziert an Springer Fachmedien Wiesbaden GmbH, ein Teil von Springer Nature 2023

Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung, die nicht ausdrücklich vom Urheberrechtsgesetz zugelassen ist, bedarf der vorherigen Zustimmung des Verlags. Das gilt insbesondere für Vervielfältigungen, Bearbeitungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.

Die Wiedergabe von allgemein beschreibenden Bezeichnungen, Marken, Unternehmensnamen etc. in diesem Werk bedeutet nicht, dass diese frei durch jedermann benutzt werden dürfen. Die Berechtigung zur Benutzung unterliegt, auch ohne gesonderten Hinweis hierzu, den Regeln des Markenrechts. Die Rechte des jeweiligen Zeicheninhabers sind zu beachten.

Der Verlag, die Autoren und die Herausgeber gehen davon aus, dass die Angaben und Informationen in diesem Werk zum Zeitpunkt der Veröffentlichung vollständig und korrekt sind. Weder der Verlag noch die Autoren oder die Herausgeber übernehmen, ausdrücklich oder implizit, Gewähr für den Inhalt des Werkes, etwaige Fehler oder Äußerungen. Der Verlag bleibt im Hinblick auf geografische Zuordnungen und Gebietsbezeichnungen in veröffentlichten Karten und Institutionsadressen neutral.

Planung/Lektorat: David Imgrund

Springer Vieweg ist ein Imprint der eingetragenen Gesellschaft Springer Fachmedien Wiesbaden GmbH und ist ein Teil von Springer Nature.

Die Anschrift der Gesellschaft ist: Abraham-Lincoln-Str. 46, 65189 Wiesbaden, Germany

Inhaltsverzeichnis

1	Einleitung	1
1.1	Warum dieses Buch?	1
1.2	Aufbau des Buches und der Buchreihe	2
2	Künstliche Intelligenz: Ein Überblick	5
2.1	Definitionsversuche: Was versteht man unter KI?	5
2.2	Eine kurze Zeitreise: Wie hat sich KI entwickelt?	8
2.3	Einordnung und Ordnungsrahmen	14
2.3.1	Einordnung	14
2.3.2	Definition und Aufbau	16
2.3.3	Teilbereiche	18
3	Machine Learning	23
3.1	Einordnung und Definition	23
3.2	Lernszenarien: Was und wie lernt ein Computer?	27
3.2.1	Terminologie	27
3.2.2	Aufgabentypen: Was soll der Computer können?	31
3.3	Lernverfahren: Ein kurzer Überblick	49
3.3.1	Grundbausteine maschineller Lernverfahren	49
3.3.2	Verfahrensklassen	51
3.4	Herausforderungen und Erfolgsfaktoren	56
3.4.1	Erfolgsfaktor Vorgehensweisen	56
3.4.2	Overfitting erkennen und vermeiden	63
3.4.3	Zentrale Schlüsselbereiche als Einflussfaktoren	68
4	Deep Learning	71
4.1	Einleitung	71
4.2	Neuronale Netze: Vom Perzeptron zum Deep Learning	72
4.2.1	Das Perzeptron	73
4.2.2	Mehrschichtige vorwärts gerichtete Netze	76
4.2.3	Verallgemeinerung des Neuronenmodells	78

4.2.4	Was und wie lernt ein neuronales Netz?	85
4.2.5	Durchführung der Back-Propagation an einem Beispiel	100
4.2.6	Ankunft beim Deep Learning – Was ist neu?	110
4.3	Deep-Learning-Architektur-Bausteine.	112
4.3.1	Feed Forward Neural Networks (FNN)	113
4.3.2	Convolutional Neural Networks (CNNs)	114
4.3.3	Recurrent Neural Networks (RNNs)	121
4.3.4	Encoder-Decoder-Architekturen	132
4.3.5	Transformer-Modelle und Attention-Mechanismen.	138
4.3.6	Lösungen aus dem Baukasten	155
4.4	Herausforderungen und Erfolgsfaktoren	157
4.4.1	Wie lässt sich der Datenhunger stillen?	157
4.4.2	Wie ist das Trainieren tiefer neuronaler Netze möglich?	158
4.4.3	Was ist erreicht und was lässt sich verbessern?	158
5	Informationsextraktion aus Texten	161
5.1	Einleitung	161
5.2	Welche Typen von Daten werden extrahiert	163
5.3	IE-Systeme und deren Entwicklungshistorie	166
5.4	Aufbau und Elemente einer IE-Pipeline	170
5.4.1	Named Entity Recognition	172
5.4.2	Coreference Resolution	176
5.4.3	Relation Extraction	178
5.4.4	Event Extraction	181
5.4.5	Named Entity und Relationship Linking	182
5.5	Informationsextraktion aus anderen Quellen	183
6	Wissensrepräsentationen	185
6.1	Ontologien	185
6.1.1	Was ist eine Ontologie?	185
6.1.2	Warum sollten Ontologien verwendet werden?	186
6.1.3	Wie können Ontologien modelliert werden?	187
6.2	Wissensgraphen	190
6.2.1	Was ist ein Wissensgraph?	190
6.2.2	Erstellung von Wissensgraphen	194
6.2.3	Anwendungsgebiete	196
	Literatur	197



1.1 Warum dieses Buch?

Sie fragen sich sicherlich, warum noch ein Buch über Künstliche Intelligenz (KI)? Das Thema Künstliche Intelligenz unterliegt aktuell einem großen Hype und KI-Technologien halten in nahezu allen erdenklichen Bereichen unseres privaten und beruflichen Alltags Einzug, sodass die Frage berechtigt ist. Was wollen und können wir mit diesem Buch anderes vermitteln, als in vielen Fachartikeln und Fachbüchern bereits steht und über viele KI-Anwendungen schon bekannt ist?

Der Titel des Buches gibt bereits darüber Aufschluss: „*Knowledge Science – Grundlagen: Methoden der Künstlichen Intelligenz für die Wissensextraktion aus Texten.*“ Es geht um die Einführung und Beschreibung von Konzepten und Methoden aus dem Bereich der Künstlichen Intelligenz, die für die Wissensextraktion aus Texten eine große Relevanz haben. Diese bilden die technische Grundlage für Anwendungen, die die Erfassung und Anwendung von Wissen jeglicher Art ermöglichen oder voraussetzen. Wir unterscheiden dabei zwischen *implizitem* und *explizitem* Wissen:

Unter explizitem Wissen versteht man Wissen, das bewusst erworben und angewandt oder bewusst gesammelt und gespeichert werden kann. Implizites Wissen ist demgegenüber unbewusstes Wissen, das man nur schwer verbalisieren kann, es ist das Ergebnis von Erfahrungen und somit kann es nur schwer an andere weitergegeben werden.¹

Der Unterschied zwischen diesen beiden Arten des Wissens lässt sich verständlich mit dem in diesem Kontext häufig zitierten Beispiel des Fahrradfahrens veranschaulichen:

¹ Diese Differenzierung zwischen explizitem und implizitem Wissen wurde mithilfe des Sprachmodells GPT-3 erzeugt, das eine Form künstlicher Intelligenz darstellt [130].

[Fahrradfahren] ist eine Fähigkeit, die sowohl implizites als auch explizites Wissen erfordert. Um Fahrrad fahren zu lernen, müssen wir zunächst einmal das explizite Wissen erwerben, also die Regeln der Fahrradbenutzung, die Bedeutung der verschiedenen Teile des Fahrrads und die Technik des Fahrradfahrens selbst. Sobald wir dieses Wissen erworben haben, können wir es dann unbewusst verarbeiten und anwenden, was dann zu implizitem Wissen führt.²

Das Ziel der in diesem Buch dargestellten Konzepte und Methoden aus dem Bereich der Künstlichen Intelligenz ist es, explizites und implizites Wissen zu identifizieren, zu extrahieren und somit verarbeitbar und nutzbar zu machen. Im zweiten Band der Buchreihe werden KI-gestützte Anwendungsfälle vorgestellt. Als Beispiel seien Assistenzsysteme genannt, die auf Basis der extrahierten Wissenssammlungen Anwenderinnen und Anwender unterstützen, indem sie das Wissen aufgaben- und bedarfsgerecht auswählen, aufbereiten und zur Verfügung stellen. Dadurch wird die Wissenssicherung und Wissensweitergabe unterstützt.

Bevor wir jedoch in spezifische Themenfelder der KI einsteigen, die für die Extraktion von Wissen aus Texten besonders relevant sind, soll für ein grundlegendes Verständnis der Thematik und der Möglichkeiten des Begriffs *Künstliche Intelligenz* eingeführt und insbesondere auch im historischen Kontext dargestellt werden.

1.2 Aufbau des Buches und der Buchreihe

Nach dieser Einleitung mit einer Darstellung der Entstehung, Entwicklung und Bestandteile der Künstlichen Intelligenz folgen Kapitel für ausgewählte Teilbereiche der KI als eine Art Werkzeugkasten für die Umsetzung von Anwendungen kognitiver Assistenten, die im zweiten Band dieser Buchreihe thematisiert werden. Dieses Kapitel kann auch als Basiswissen für die Leser herangezogen werden, die einen ersten Überblick über die Bereiche Machine Learning und Deep Learning erlangen wollen und kann dementsprechend auch als eigenständiges Werk gelesen werden. Die folgenden Kapitel zur Informationsextraktion aus Text und Wissensrepräsentation auf Basis von Knowledge Graphen greift den starken Fokus auf Wissensextraktion und Wissensbewahrung mithilfe von Verfahren des NLP auf.

Im zweiten Band der Buchreihe werden zunächst Grundlagen aus der Perspektive der Anwendung dargestellt. Es geht dabei einerseits um die Rolle des Wissens insbesondere bei Arbeitsabläufen, und wie der Mensch bei wissensintensiver Arbeit unterstützt werden kann und ganz allgemein darum, wie sinnvoll mit der Ressource Wissen im Unternehmen umzugehen ist, also um Wissensmanagement. Wir erläutern daher zunächst, welche Aufgaben sich automatisieren oder unterstützen lassen und beleuchten die Rolle und Gestaltung von Assistenzsystemen am Arbeitsplatz der Zukunft. Anschließend gehen wir konkreter auf kognitive Assistenzsysteme, sinnvolle Unterscheidungsmerkmale und den aktuellen Stand der Technik ein. Eine kompakte Einführung in das Wissensmanagement rundet die Grundla-

² Auch diese Erklärung stammt von GPT-3 [130] – aber keine Sorge, der weitaus größere Teil des Buches wurde von Menschen geschrieben.

gen für die Anwendungsfälle ab. Der Hauptteil des zweiten Bandes greift die KI-Grundlagen aus diesem Band sowie die anwendungsorientierten Grundlagen auf und setzt diese in den Anwendungsbezug. Es werden Anwendungen aus unterschiedlichen Bereichen verständlich erläutert. Zum Abschluss des zweiten Bandes werden die Anwendungen mittels eines Perspektivenkonzepts eingeordnet und übersichtlich zusammengefasst.



Zusammenfassung

Wir haben den Begriff *Künstliche Intelligenz (KI)* bereits mehrfach verwendet, ohne ihn explizit zu definieren. Was genau ist Künstliche Intelligenz? Gibt es eine exakte und allgemeingültige Definition? Seit wann gibt es den Begriff und wie hat sich das Thema als Fachgebiet über die Zeit entwickelt? Die folgenden Abschnitte sollen diese Fragen beantworten. Anschließend nehmen wir eine eher technische Perspektive ein, stellen Teilbereiche der KI vor und ordnen diese sie ein.

2.1 Definitionsversuche: Was versteht man unter KI?

Um zu verstehen, was *Künstliche Intelligenz* bedeutet, erscheint es sinnvoll, zunächst die Bedeutung des Attributs *künstlich* zu erläutern. Es bedeutet in diesem Kontext *nicht natürlich*, also *technisch* oder vom Menschen geschaffen. So wird Künstliche Intelligenz gelegentlich auch als rechnergestützte Intelligenz oder im Englischen auch als *Computational* oder *Machine Intelligence* bezeichnet und deutlich von der natürlichen und somit insbesondere menschlichen Intelligenz abgegrenzt.

Was aber genau ist *Intelligenz*? Es zeigt sich, dass eine allgemeingültige Definition aufgrund der vielschichtigen Facetten und Ausprägungen, die Intelligenz hat, nicht einfach zu finden ist.

Eine pragmatische Herangehensweise ist die Charakterisierung von Intelligenz mittels einer offenen (nicht vollständigen) Ansammlung verschiedener Eigenschaften in Anlehnung an menschliches Verhalten als eine Art Wunschliste für künstliche (rechnergestützte) *intelligente Systeme*: [74]

- **Wahrnehmung:** den Zustand der Umwelt über Sensoren aufnehmen (Rohdaten), daraus Informationen ableiten, wie die Erkennung bestimmter Objekte, die abgeleiteten Informationen mit bestehenden Informationen zusammenführen und interpretieren,
- **Schlussfolgern:** auf Basis der verfügbaren Informationen auch unter Unsicherheit und unvollständigen Informationen deduktiv, induktiv oder transduktiv Schlüsse ziehen,
- **Handeln:** zur Erkundung und Veränderung der Umwelt auf Basis der verfügbaren Informationen kontrollierte Aktionen ableiten und einsetzen sowie Werkzeuge entwickeln und nutzen,
- **Planung und zielgerichtete Problemlösung:** eine Abfolge von Handlungsschritten zur Lösung eines Problems ermitteln und festlegen,
- **Kommunikation:** mit anderen intelligenten Systemen (einschließlich Menschen) kommunizieren,
- **Anpassungsfähigkeit und Lernen:** zum Umgang mit neuen oder veränderten Situationen auf Basis von gewonnenen Informationen und Erfahrungen das Verhalten anpassen oder neue Verhaltensweisen lernen,
- **Autonomie:** sich eigene Ziele setzen und selbst über den Weg zur Zielerreichung entscheiden,
- **Kreativität:** neue Wege zur Problemlösung erkunden,
- **Reflexion und Bewusstheit:** eigene interne Prozesse, Ziele und Entscheidungen sowie die anderer reflektieren,
- **Ästhetik:** beim Entscheiden und Handeln ästhetische Prinzipien berücksichtigen,
- **Organisation:** mit anderen intelligenten Systemen interagieren und sich mit ihnen abstimmen.

Während die meisten Menschen als natürliche intelligente Systeme diese Eigenschaften mehr oder weniger ausgeprägt aufweisen, sind künstliche intelligente Systeme noch weit davon entfernt. Sie erreichen oft nur Teilmengen dieser Eigenschaften, und zwar umso größere, je stärker die Anwendungsdomäne mit ihren Problemstellungen eingeschränkt ist [74]. Je enger eine Aufgabe definiert ist, desto erfolgreicher können künstliche intelligente Systeme agieren.

Die ersten drei Eigenschaften, also Wahrnehmung, Schlussfolgern und Handeln, hat Patrick Winston, amerikanischer Informatiker und langjähriger Leiter des AI-Labs am Massachusetts Institute of Technology (MIT), schon früh aufgegriffen und Künstliche Intelligenz als wissenschaftliche Disziplin definiert, die Berechnungsverfahren entwickelt und untersucht, die es einem System ermöglichen, wahrzunehmen, zu schlussfolgern und zu handeln [191]. Ziel dieser Berechnungsverfahren ist es also, die menschliche Wahrnehmungs- und Verstandesleistung abzubilden, indem Computerprogramme entwickelt werden, die die Fähigkeit haben, Problemlösungsbereiche zu bearbeiten, die bisher nur vom Menschen gelöst werden konnten [142].

Da es zurzeit aber noch kein KI-System gibt, das generalistisch agiert, werden Lösungen Schritt für Schritt entwickelt und so der Fokus jeweils auf einzelne, sehr konkrete Probleme gelegt. Dies führt zur Unterscheidung zwischen *starker* und *schwacher KI*.

Unter einer *schwachen KI* wird ein System verstanden, das ein oder wenige sehr konkret eingegrenzte Anwendungsprobleme lösen kann, um eine vermeintlich menschliche Intelligenz zu simulieren oder zu imitieren. Diese KI-Systeme sind in der Domäne, für die sie entwickelt wurden, meist extrem leistungsfähig, in anderen Domänen allerdings kaum einsetzbar. Es fehlt die Eigenschaft, bekannte Lösungsmuster auf neue Probleme zu übertragen—eine Eigenschaft, die gerade auch die menschliche Intelligenz auszeichnet.

Eine *starke KI* ist ein System, das die oben aufgeführte umfangreiche Liste an Eigenschaften abdeckt und damit menschliche Problemlösungskreativität, Selbstbewusstsein und sogar Emotionen abbilden kann. Sie ist ein System, das generisch Probleme lösen kann, ohne vorher dafür programmiert oder trainiert worden zu sein. Auch der Transfer von vorhandenem Wissen und Fähigkeiten auf eine andere Domäne stellt für ein System mit starker KI keine Schwierigkeiten dar. Ein derart autark agierendes System würde uns Menschen in nichts nachstehen und sich selbst ständig weiterentwickeln. Es würde am Ende zur sogenannten technologischen Singularität kommen. Dies wäre der Moment, in dem die Intelligenz der KI diejenige des Menschen übertrifft, was dann auch als *Super-Intelligenz* bezeichnet wird. Ohne entsprechenden Willen und die Fähigkeit zur Erklärung getroffener Entscheidungen und Handlungen, könnte es für einen Menschen immer schwieriger werden, diese nachzuvollziehen. Ob die Ziele der Super-Intelligenz für Menschen rational verständlich sind, wissen wir nicht. Durch die genannten Entwicklungsmöglichkeiten werden daher nicht nur technische, sondern auch ethische Fragen aufgeworfen. Bislang existieren allerdings keine starken KI-Systeme und auch in naher Zukunft ist nicht davon auszugehen, dass eine starke KI mit generischer Problemlösungsfähigkeit geschaffen wird.

Sehr wohl gibt es aber bereits KI-Systeme, die in einzelnen Spezialgebieten die menschliche Geistesleistung deutlich übertreffen, wie bei dem Strategiespiel *Go*. Im April 2022 hat Google sogar gezeigt, dass KI-Systeme mit sogenannten *Large Language Models* in der Lage sind, den Menschen in etlichen Bereichen der Sprachverarbeitung zu übertreffen. Die Fähigkeiten des in diesem Zusammenhang entwickelten *Pathways Language Model (PaLM)* mit unglaublichen 540 Mrd. Parametern hat Google mittels des BIG-bench-Datensatzes getestet [27].¹ Die Forscherinnen und Forscher von Google wählten 150 Tasks aus und evaluierten das Sprachmodell im Vergleich zur Leistung von Menschen. Im Ergebnis übertrifft das Modell die durchschnittliche menschliche Leistung in 65 % der getesteten Tasks [126]. Auch wenn es aufgrund der beeindruckenden Ergebnisse scheint, als wäre hier eine starke KI entwickelt worden, handelt es sich doch weiterhin nur um eine Form schwacher KI.

¹ BIG-bench steht für Beyond the Imitation Game Benchmark. Dabei handelt es sich um 209 Tasks, die von arithmetischem Verständnis über die Beantwortung von Multiple-Choice-Fragen, über Gender-Fairness-Test für Sprachmodelle bis zu logischen Ableitungen eine große Bandbreite von verschiedenen Tests abbilden [169].

Trotz der Unfähigkeit, generisch Probleme lösen zu können, zeigen diese imponierenden Modelle auf, wo die Reise hingeht. Die Entwicklung derartiger Sprachmodelle bringt für unseren Themenkontext *Knowledge Science* einen enormen Mehrwert in der automatisierten Wissensextraktion und -verwendung, da sie die dafür notwendigen Aufgaben der Sprachverarbeitung (Natural Language Processing) auf ein vorher unerreichtes Niveau anheben.

2.2 Eine kurze Zeitreise: Wie hat sich KI entwickelt?

Künstliche Intelligenz bzw. die Überlegung, wie durch Maschinen das menschliche Denken nachempfunden werden kann, ist keine Erscheinung der letzten Jahre, sondern reicht mehrere Jahrhunderte zurück. Honavar beginnt seine Chronologie der Künstlichen Intelligenz bereits im Jahre 384 vor Christus unter Bezugnahme auf Aristoteles, der zwischen Materie und Form unterscheidet und damit die Grundlagen für die Abstraktion zwischen dem Untersuchungsgegenstand und dessen Darstellung legt. Dies bildet die Grundlage unserer heutigen modernen Informatik [74]. 350 vor Christus entwickelte dann Panini eine formale Grammatik des Sanskrit², was wiederum im Jahre 1956 von Noam Chomsky als Grundstein für die Theorie der syntaktischen Strukturen aufgegriffen wurde [74].

Viele weitere Meilensteine wären zu nennen, aber in dieser Arbeit soll der Fokus auf der Entwicklung seit den 50er-Jahren liegen. Beginnend mit der Formulierung des sogenannten Turing-Tests durch Alan Turing im Jahr 1950, der eine Idee zur Feststellung formulierte, ob ein Computer ein dem Menschen gleichwertiges Denkvermögen besitzt. Der Test, von ihm selbst als *Imitation Game* bezeichnet, soll die Frage beantworten, ob Computer denken können: „Can machines think?“ Der Test wurde von Turing zunächst nur theoretisch skizziert und erst nach seinem Tod konkreter ausformuliert.

Der Test läuft als Spiel mit drei Personen ab, einer Frau (A), einem Mann (B) und einem Interrogator (C). Der Interrogator ist räumlich von A und B getrennt. Ziel des Spiels ist es, dass C durch Befragung der Personen A und B herausfindet, welche der beiden die Frau und wer der Mann ist. A hat dabei die Aufgabe, die Antworten an C so zu formulieren, dass die Identifikation schwierig ist. B hingegen hat die Aufgabe, dem Interrogator C durch klare und präzise Antworten zu helfen. Beispielsweise könnte B antworten: „Ich bin ein Mann, bitte höre nicht auf A.“ Damit nicht anhand der Stimme erkannt werden kann, um wen es sich handelt, wird die Befragung ausschließlich per Text z.B. über ein Chatprogramm vorgenommen. Um das Spiel als Test für die Denkleistung eines Computers zu verwenden, stellte Turing die Fragen: „Was würde passieren, wenn eine Maschine die Aufgabe von A übernehmen würde?“ Und würde C genauso oft falsche oder richtige Entscheidungen treffen, wie wenn dieses Spiel ausschließlich mit Menschen gespielt werden würde? Wäre dies der Fall, würde man annehmen, dass die Maschine im Verständnis und in der Beantwortung der menschlichen Denkleistung gleichwertig sei [176].

² Sanskrit ist eine Bezeichnung für die verschiedenen Varietäten des Altindischen.

Der Turing-Test ist auch heute noch gültig, wenn auch mit einigen Modifikationen. Zudem werden Stimmen laut, die fordern, den Test durch andere Tests zu ersetzen, da die großen Sprachmodelle wie GPT-3 und PaLM in ihrer Leistung so weit fortgeschritten seien, dass der originäre Turing-Test nicht mehr vollständig Bestand habe.

Ein weiteres wichtiges Ereignis in der Geschichte der Künstlichen Intelligenz ist die Konferenz mit dem Titel „Darthmouth Summer Research Project on Artificial Intelligence“, die im Sommer 1956 am Darthmouth College in Hanover, New Hampshire, stattfand. Auf dieser Konferenz verwendeten die Wissenschaftler John McCarthy, Marvin Minsky, Nathaniel Rochester und Claude Shannon als Antragsteller zusammen mit den weiteren Teilnehmern Ray Solomonoff, Oliver Selfridge, Allen Newell, Herbert Simon, Trenchard More und Arthur Samuel erstmalig den Begriff „Artificial Intelligence“ und markierten somit die Geburtsstunde der Künstlichen Intelligenz als akademisches Feld. Die Konferenz war als zweimonatiger Workshop angesetzt, bei dem sich die Wissenschaftler über das neue Themengebiet austauschen und funktionierende Lösungen entwickeln sollten. Insbesondere waren in dem Antrag für die Konferenz die folgenden Themengebiete aufgeführt: [114]

- Automatische Computer
- Wie können Computer programmiert werden, um Sprache zu verwenden?
- Neuronale Netze
- Theorie über die Größe einer Berechnung
- Selbstoptimierung von Algorithmen
- Abstraktionen
- Zufälligkeit und Kreativität

Die Verwendung des Begriffs Künstliche Intelligenz für das neue Forschungsgebiet war dabei unter den Teilnehmern keineswegs unstrittig. Es standen auch Vorschläge wie *komplexe Informationsverarbeitung* oder *Automatenstudie* im Raum. Ganz klar war aber allen – und das zeigte sich auch schon im Antragstext – dass es nicht um die Schöpfung einer künstlichen Intelligenz gehen sollte, sondern darum, eine Maschine dahin gehend zu programmieren oder anzupassen, dass sie Intelligenz bzw. die menschliche Denkleistung simuliert [135].

Ein weiterer größerer Meilenstein war die Entwicklung des ersten Chatbots durch den deutsch-amerikanischen Informatiker sowie Wissenschafts- und Gesellschaftskritiker Joseph Weizenbaum [189] im Jahre 1964. In den Jahren 1964 bis 1966 entwickelte Weizenbaum ELIZA, eines der ersten Programme zur automatischen Sprachverarbeitung (Natural Language Processing), das den Zweck hatte, eine Maschine ein Gespräch mit einem Menschen führen zu lassen, wobei das Programm einen Psychotherapeuten imitierte. ELIZA ist ein regelbasierter Chatbot, der tatsächlich die Illusion einer intelligenten Maschine und damit auch einer natürlichen Konversation erzeugte. Dabei kamen lediglich Schlüsselworterkennung sowie geschickte Anwendung von Regeln zur Umformung von Sätzen zum Einsatz. Dennoch waren die Ergebnisse insbesondere für die damalige Zeit beeindruckend

gut. Daher wies Weizenbaum immer wieder darauf hin, dass hier lediglich eine Illusion entsteht und die Maschine selbst nicht wirklich intelligent sei [187].

Auch in einem im Jahre 1966 veröffentlichten Artikel äußerte sich Weizenbaum zu dieser Illusion wie folgt [182]:

For in those realms machines are made to behave in wondrous ways, often sufficient to dazzle even the most experienced observer. But once a particular program is unmasked, once its inner workings are explained in language sufficiently plain to induce understanding, its magic crumbles away; it stands revealed as a mere collection of procedures, each quite comprehensible.

Dennoch fand hier ein Durchbruch für die Künstliche Intelligenz und insbesondere für das Natural Language Processing statt.

Auch erwähnenswert ist die Entwicklung der *Backpropagation* im Jahre 1970. Backpropagation ist eine Methode, mit der neuronale Netze aus ihren Fehlern lernen können. Erstmals wurde dieses Verfahren 1970 vom finnischen Studenten Seppo Linnainmaa in seiner Masterarbeit und 1976 in einem Artikel veröffentlicht [105]. Die Arbeit von Linnainmaa basierte auf einer Idee von Henry J. Kelley, die dieser 1960 in seinem „Gradient Theory of Optimal Flight Paths“ formulierte [87]. Erst 30 Jahre später wurde das Prinzip für die Anpassung der Gewichte in neuronalen Netze durch Rummelhart, Hinton und Willms in ihrem Artikel „Learning representations by back-propagating errors“ herausgearbeitet und damit populär gemacht.

Die Grundidee dabei ist es, nach regulärem Durchleiten von Eingangssignalen durch ein neuronales Netz (forward pass) auf Basis von Gradienten einer Fehlerfunktion Signale zur Anpassung der Gewichte rückwärts durch das neuronale Netz zu leiten (daher der Name Backpropagation), um so iterativ den Prognosefehler zu minimieren [19]. Hierbei wird mittels der Trainingsdaten durch eine vorwärts gerichtete Berechnung zunächst anhand der vorhandenen Netzgewichtungen ein Output berechnet, der mit dem wahren Wert verglichen wird, sodass der Fehler ermittelt werden kann. Im nächsten Schritt werden dann durch die sogenannte Backpropagation die notwendigen Anpassungen der Gewichte des Netzes vorgenommen, die den Fehler minimieren. Dies wird so lange wiederholt, bis ein (lokales) Minimum der Fehlerfunktion erreicht ist. Durch dieses Vorgehen kann das Netz lernen [151].

Im Jahr 1972 wurde an der Stanford Universität erstmals mit einer regelbasierten Logik und der Programmiersprache LISP ein Expertensystem entwickelt, das auf Basis einer Wissensbasis die Anwenderinnen und Anwender bei Treffen von Entscheidungen unterstützen konnte: Das System MYCIN umfasste ungefähr 600 Regeln und unterstützte bei der Erkennung von Infektionen und gab Empfehlungen für die passende Medikamentendosis angepasst an das Körpergewicht des Patienten [175].

Die meisten KI-basierten Lösungen, die zwar oft prinzipiell funktionierten, scheiterten jedoch meist in echten Anwendungen. Ein wesentlicher Grund dafür war die zu der Zeit nur in sehr begrenztem Umfang verfügbare Rechenleistung. Viele Ideen konnten nur sehr einfache, beinahe triviale Probleme lösen, aber keine komplexen, anwendungsnahen Probleme. Hier waren die Visionen größer als die realen Möglichkeiten. Interessant ist, dass viele Ideen

und Lösungsvorschläge aus dieser Zeit in den letzten zehn bis 15 Jahren insbesondere mit Verfügbarkeit entsprechender Rechenleistung erfolgreich zur Anwendung gekommen sind [34].

Doch auch eine Winterzeit endet einmal. Die erste Winterzeit dauerte etwa bis 1980, bis sie langsam durch die nächste Boom-Welle abgelöst wurde. Auslöser der nur kurz andauernden zweiten Welle waren die Entwicklung von meist regelbasierten Expertensystemen sowie die stärkere Fokussierung auf Anwendungen und kommerzielle Nutzung [161]. In dieser Phase entwickelte z. B. Lee Erman et al. das Blackboard Design Pattern, das verwendet werden kann, um eine Wissensbasis (Blackboard) aufzubauen und die in dieser Wissensbasis gespeicherten Quelldaten iterativ zu aktualisieren, wenn neue Probleme bzw. Anfragen auftreten, bis eine Lösung gefunden wurde [175].

Ausschlaggebend für die Expertensysteme dieser Epoche war die Tatsache, dass die definierten Regeln, meistens Wenn-Dann-Regeln, mittels eines Top-Down-Approach manuell ermittelt wurden. Es wurde angenommen, dass das Abbilden von Domänenwissen durch die Befragung von Domainexperten mittels Regeln der beste Weg sei, eine KI aufzubauen. 1984 sprangen auch Zeitschriften und Zeitungen wieder auf den Hype auf. So lautete eine Titelschlagzeile der *Business Week* „AI: It’s Here“ [161].

Aufgrund des Hypes fürchteten einige Wissenschaftlerinnen und Wissenschaftler, dass die Erwartungen erneut nicht erfüllt werden und es zu einem zweiten Winter kommen könnte. So gab es auf der AAAI-Konferenz ein Panel mit dem Titel „The Dark Ages of AI – Can we avoid or survive them?“, auf dem diskutiert wurde, wie man einen zweiten Winter abwenden könnte. Die größte Sorge war natürlich, dass Forschungsgelder gekürzt und man dadurch ausgebremst werden würde.

Wie in Abb. 2.1 ersichtlich, flachte die zweite Boom-Zeit dann schon in den 90er-Jahren wieder ab. Hintergrund hierfür war, dass die hochgepriesenen Expertensysteme etliche Beschränkungen aufwiesen. John McCarthy kritisierte beispielsweise, dass die Systeme

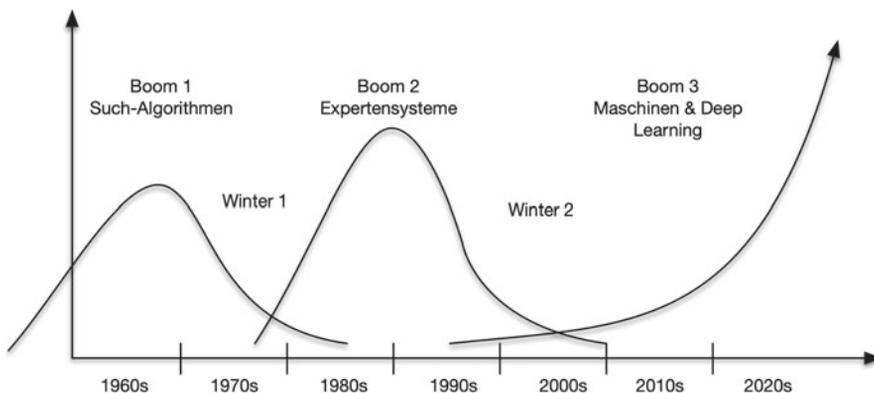


Abb. 2.1 Winter- und Boom-Zeiten Künstlicher Intelligenz. (In Anlehnung an [60])

keinen gesunden Menschenverstand nachbilden konnten und dass sie sich nicht ihrer eigenen Grenzen bewusst waren. Andere Kritikpunkte waren, dass viele Themen so komplex waren, dass die Entwicklerinnen und Entwickler Schwierigkeiten hatten, sie in Wenn-Dann-Regeln abzubilden. Vor allem Computer Vision und Sprachverarbeitung waren so komplex, dass sie kaum vollständig in Regelsystemen abgebildet werden konnten. Auch hier zeigte sich, dass die Unzulänglichkeiten, die dann sichtbar wurden, wieder zu einem verringerten Interesse in der Bereitstellung von Forschungsgeldern führte und somit den zweiten KI-Winter einleiteten [161].

Das Ende dieser Periode sollte Anfang der 2000er-Jahre durch einen Paradigmenwechsel weg von regelbasierten zu statistischen und maschinellen Lernverfahren eingeleitet werden. Gerade in den Anfängen der 2000er-Jahre wurden viele Lernverfahren wie Support-Vektor-Maschinen, Bayes'sche Netze, Markov-Modelle und vor allem auch neuronale Netze eingesetzt. Dieser Paradigmenwechsel in der KI führte dazu, dass eine länger andauernde Phase eines KI-Frühlings begann, die bis heute anhält. Die Anwendung maschineller Lernverfahren zeigte an vielen Stellen, dass diese praktikabler waren als die regelbasierten Ansätze. Ein Schlüsselereignis dabei war die von der DARPA ausgerufene Grand Challenge zum autonomen Fahren im Jahre 2005. Dotiert mit zwei Millionen Dollar Preisgeld, bestand die Aufgabe darin, autonome Fahrzeuge zu entwickeln, die eine 132 Meilen lange Strecke durch die Wüste von Nevada zurücklegen konnten. Gewinner des Wettbewerbs war der deutsche Wissenschaftler Sebastian Thrun mit seinem Team der Stanford University. Die New York Times griff dieses Ereignis nach den Rückschlägen der Vorjahre auf und titelte, dass hinter der künstlichen Intelligenz ein Geschwader von klugen, echten Menschen stehe [112]. Und diese äußerten, dass mit den jüngsten Erfolgen nun ein KI-Frühling begonnen habe [112].

Mit der starken Zunahme der Nutzung maschineller Lernverfahren bei Entwicklung von KI-Lösungen startete auch die Deep-Learning-Revolution. Angetrieben durch Cloud-Computing und Digitalisierungsbestrebungen sowie natürlich auch der zunehmenden Erzeugung und Bereitstellung von Daten durch IoT-Geräte, standen nun einer deutlich breiteren Gruppe an KI-Forscherinnen und Forschern mehr Ressourcen zur Verfügung [161]. Dies führte etwa im Jahre 2021 dazu, dass etliche Erfolge im Bereich des maschinellen Lernens gefeiert werden konnten.

Neuronale Netze zeigten nun langsam ihre Überlegenheit in vielen Anwendungsbereichen. So gewann erstmalig ein neuronales Netz die jährlich stattfindende ImageNet Large Scale Visual Recognition Challenge, bei der es darum geht, den Inhalt von Bildern zu erkennen und damit eine automatische Annotation der Elemente in den Bildern vornehmen zu können [91]. Aber auch beispielsweise in der Sprachverarbeitung machten KI-Systeme bedeutende Fortschritte. So gewann IBMs *Watson*-System in der Quizshow *Jeopardy*. Apple veröffentlichte im Jahr 2011 *Siri*, Google startete 2012 den Sprachdienst *Google Now*, Microsofts *Cortana* gibt es seit 2013 und Amazon veröffentlichte im Jahre 2014 den virtuellen Sprachassistenten *Alexa* [125].

Wie schon erwähnt, sind zwei wesentliche Faktoren dafür verantwortlich, dass die Deep-Learning-Revolution so durchstarten konnte. Erstens die enorme Menge an oft frei ver-

fürbaren Daten für viele Anwendungsbereiche und zweitens die Möglichkeit, über Cloud-Anbieter auf Rechenleistung etwa in Form von Grafikkarten zugreifen zu können. Die große Bedeutung der Daten gerade im Vergleich zum Einfluss der Lernverfahren wird sehr prägnant im Artikel mit dem Titel „The Unreasonable Effectiveness of Data“ zum Ausdruck gebracht: Ein einfacher Algorithmus, gefüttert mit mehr Daten, wird immer einen komplexeren Algorithmus schlagen, der mit weniger Daten gefüttert wurde [61, 161].

Die positive inhaltliche Entwicklung wird zusätzlich durch immer großzügigere Förderungsgelder durch öffentliche und private Forschungsinitiativen angetrieben. So wurden im Jahre 2011 670 Mio. US\$ pro Jahr in KI-Start-Ups investiert. Im Jahr 2021 belief sich die Summe dann schon auf 36 Mrd. US\$ und im Jahr 2022 wurden im ersten Halbjahr bereits 28 Mrd. US\$ als Fremdkapital für Start-Ups bereitgestellt [171].

Die Darstellung der historischen Entwicklung soll mit einer Überlegung abschließen, welches die ausschlaggebenden Faktoren dafür sind, dass ein KI-Winter eintritt, und ob wir gerade auf einen neuen KI-Winter zusteuern. Schumacher zeigte in einer Ausarbeitung im Jahre 2019, dass es primär die drei folgenden Kriterien waren, die zu den bisherigen KI-Winter-Phasen geführt haben: [161]

- Geschürte Erwartungen und Versprechungen entsprechen nicht der Realität,
- Forschungsgelder werden reduziert,
- technologische Barrieren bremsen oder blockieren Entwicklungen.

Bei Betrachtung des erstens Punktes ist zu konstatieren, dass tatsächlich gerade ein enormer Hype zum Thema KI entstanden ist. Es scheint fast so, als müsse alles mit KI angereichert werden, um am Markt bestehen zu können. Dieser durch den Markt und die Wirtschaft getriebene Hype birgt ein enormes Risiko in Bezug auf übertriebene Erwartungen und Versprechen.

Demgegenüber zeigt sich aktuell allerdings kein wesentlicher Rückgang an Forschungsgeldern. Auch muss erwähnt werden, dass in der aktuellen Situation die Unterstützung von Forschungsprojekten nicht nur aus staatlich initiierten Forschungsaktivitäten entstanden ist, sondern dass vielmehr viele Unternehmen ebenfalls in KI-Forschung investieren und somit eine stabilere Finanzierung vieler Ideen und Aktivitäten existiert. Dies zeigt sich laut Schumacher auch an der Marktvorschau bis zum Jahr 2025, in der Forschungsausgaben zwischen 126 Mrd. und 644 Mrd. US\$ prognostiziert werden [161].

Im dritten und letzten Punkt erkennen wir aktuell jedoch kaum signifikante Hindernisse. Insbesondere die prominenten Sprachmodelle wie ChatGPT entfalten ein beeindruckendes Potenzial. Auch für die in dieser Buchreihe thematisierten Anwendungen sind speziell die jüngsten Fortschritte im Bereich der Sprachverarbeitung hervorzuheben. Diese basieren vorrangig auf Transformer-Modellen, zu denen auch die GPT-Modell-Familie von OpenAI zählt. Aber auch in anderen Bereichen, wie die Entwicklung von Generative Adversarial Networks (GANs), kam es zu beeindruckenden Fortschritten. Diese Netzwerke können Bilder und Audio-Dateien zu erzeugen, die kaum von der Realität zu unterscheiden sind.

Aber es sind nicht nur positive Stimmen in Bezug auf die Entwicklung zu hören. So kritisiert Francois Collet, Entwickler des Deep-Learning-Frameworks *Keras* und damit eine bedeutende Stimme in der Community, dass zwar nie mehr Menschen an Deep Learning gearbeitet hätten als heute, es aber dennoch die langsamste Entwicklungsrate in den vergangenen fünf Jahren gebe [26, 161].

Zusammenfassend lässt sich festhalten, dass die Entwicklung in der KI noch stabil ist und zudem genügend Gelder im Markt vorhanden sind, sodass es kurzfristig nicht zu einem weiteren KI-Winter kommen sollte. Das bedeutet aber auch, dass mehr Fortschritte insbesondere für die wirtschaftliche Nutzung von KI geschaffen werden müssen und dass die Erwartungen und Versprechen leicht reduziert werden müssen. Anderenfalls erleben wir womöglich in den nächsten Jahren erneut eine Winter-Phase mit stark abgeschwächtem Fokus auf KI.

2.3 Einordnung und Ordnungsrahmen

Dieser Abschnitt soll einen Überblick über die verschiedenen anwendungsorientierten und grundlegenden Teilbereiche der Künstlichen Intelligenz (KI) geben. Nach einer Einordnung und Charakterisierung der KI als Fachgebiet werden die verschiedenen Bereiche motiviert und im Gesamtkontext eingeordnet. Ausgewählte Bereiche, die als Grundlagen für die in diesem Buch betrachteten Anwendungsfälle im Kontext des *Knowledge Science* besonders relevant sind, werden in den folgenden Abschnitten detaillierter betrachtet.

2.3.1 Einordnung

Wie wir soeben in der Zeitreise durch die Entstehung und Entwicklung der Künstlichen Intelligenz (KI) festgestellt haben, gibt es bislang keine einheitliche Definition für den Begriff. Je nach Perspektive werden unterschiedliche Aspekte beispielsweise bezüglich der zu erreichenden Anwendungen und Ziele oder der Art und Weise, wie diese erreicht werden sollen, in den Vordergrund gerückt.

Die meisten der KI-Definitionen haben jedoch gemeinsam, dass ein System – virtuell als Software oder auch physisch in Verbindung mit Hardware – Aufgaben lösen soll, die eine gewisse Form von Intelligenz erfordern, wollte ein Mensch sie ausführen. Das Problem dabei ist natürlich, dass sich *Intelligenz* mit ihren vielen Facetten nur schwer greifen lässt. Die KI gilt dabei als Teilbereich der Informatik und der Fokus liegt letztlich auf der Automatisierung intelligenter (menschlicher) Entscheidungsstrukturen.

Es fällt auf, dass in dieser Perspektive der Mensch mit seiner Intelligenz als Vorgabe oder gar Vorbild eine sehr zentrale Rolle spielt: KI-Systeme sollen Aufgaben so lösen, wie der Mensch es täte. Das impliziert eine gewisse Form von Eigenständigkeit und Anpassungsfähigkeit an unbekanntere Situationen. Wenn einmal erworbene Fähigkeiten für eine Lösung

nicht mehr ausreichen, muss das Verhalten entsprechend angepasst werden, d. h. das System muss lernfähig sein.

Wenn ein KI-System menschliche Entscheidungsstrukturen unterstützen oder automatisieren soll, dann sind Analogien zum menschlichen Vorgehen und eine Orientierung daran naheliegend und in gewissem Rahmen auch hilfreich. Warum aber sollte verlangt werden, dass der Computer dabei teilweise menschliche Züge annimmt und idealerweise nicht von einem Menschen zu unterscheiden sei, wie beim Turing Test oder genauer Imitation Game, wie Turing selbst den Test bezeichnet hatte?

Wenngleich diese Forderung in ausgewählten Anwendungsbereichen wie der Unterstützung bei der Betreuung pflegebedürftiger Menschen eine wichtige Rolle für die Akzeptanz spielen kann, so ist es doch in den meisten anderen Fällen unerheblich, ob sich KI-Systeme „menschlich“ verhalten. Vielmehr scheint es grotesk, dass ein technisches System unbedingt menschliche Züge annehmen müsse, damit es für die Anwendung infrage kommt. Schließlich muss ein Flugzeug auch nicht aussehen wie ein Vogel oder sogar von Vögeln als ihresgleichen akzeptiert werden. Es muss einfach nur spezifische Eigenschaften haben, die das Fliegen ermöglichen.

Deshalb werden wir in der folgenden Betrachtung nicht voraussetzen, dass KI-Systeme zwingend menschliche Züge annehmen. Lediglich die Aufgaben selbst und die Inspiration für die Lösung sind essenziell für die Entwicklung intelligenter Systeme und daher für auch für unsere Betrachtung. Schließlich wird oft nur verlangt, dass ein KI-System rational im Rahmen des Kontexts mit aktuellen Gegebenheiten und Möglichkeiten sowie eines Kriteriums zur Bewertung des Erfolgs die besten Aktionen für die Zielerreichung auswählt und ausführt [153].

Wenn Aufgaben und Bewertungskriterien durch den Menschen vorgegeben sind, dann sprechen wir von *schwacher KI*. Genau genommen wird bei dieser Form intelligentes Verhalten lediglich *imitiert*. *Starke* oder *generelle KI* soll dagegen „echte“ Intelligenz entwickeln, die die menschliche Intelligenz sogar übertreffen kann. Diese Systeme sind dann nicht mehr auf vorgegebene Aufgaben beschränkt, sondern können theoretisch aus eigenem Antrieb handeln und auch neue Aufgaben lösen. Als wichtige Bausteine für eine starke KI werden Aspekte wie Emotionen, Selbsterkenntnis und letztlich Bewusstsein gezählt, insbesondere auch, um sich aus eigenem Antrieb eigene Ziele setzen zu können, die nicht durch den Menschen vorgegeben sind. Aktuell werden diese Fähigkeiten noch nicht erreicht.

Unser Fokus liegt auf der Lösung ausgewählter Aufgaben. Zu beachten ist allerdings, dass diese Aufgaben immer komplexer werden. Komplexe Aufgaben lassen sich nach dem Teile-und-Herrsche-Prinzip oft besser oder zumindest einfacher lösen, wobei sie in weniger komplexe Teilaufgaben heruntergebrochen werden. Folglich werden einzelne Aufgaben nicht mehr nur isoliert betrachtet. Vielmehr müssen ihre Teilaufgaben geeignet orchestriert und ihre Lösungen verknüpft werden, um komplexere Aufgaben zu lösen.