
lernOS KI Leitfaden

KI wird den Menschen nicht ersetzen - aber der Mensch mit KI wird den Menschen ohne KI ersetzen

Version 0.x (10.08.2023)

Inhaltsverzeichnis

1	Über lernOS	3
2	Grundlagen	4
2.1	Grundidee der lernOS Leitfäden	4
2.2	Geschichte der Künstlichen Intelligenz	5
2.2.1	ToDo	7
2.3	## Künstliche Intelligenz und Maschinelles Lernen	7
2.4	Neuronale Netzwerke	14
2.5	KI-Anwendungsfelder	14
2.6	KI Modelle	15
2.6.1	Wieso sprechen wir von Modellen?	15
2.6.2	Überblick zu Modellen generativer KI	16
2.6.3	Hinweis	18
2.7	KI-Tools und -Dienste	18
2.7.1	Tipps & Empfehlungen	19
2.7.2	KI-gestützte Multi-Tools	19
2.7.3	KI-gestützte Text-Tools	20
2.7.4	KI-gestützte Codierungshilfen	21
2.7.5	KI-gestützte Bild-Tools	21
2.7.6	KI-gestützte Audio-Tools	22
2.7.7	KI-gestützte Video-Tools	22
2.8	## KI und Gesellschaft: eine Reflexion über Implikationen und Verantwortung	23
2.9	Weiterführende Informationen und Links	26
3	Lernpfad	27
4	Lernpfad	28
4.1	Woche 0	28
4.2	Woche 1	28
4.3	Woche 2	29
4.4	Woche 3	29
4.5	Woche 4	30
4.6	Woche 5	30
4.7	Woche 6	30
4.8	Woche 7	30
4.9	Woche 8	30
4.10	Woche 9	30

4.11	Woche 10	30
4.12	Woche 11	30
5	Anhang	30
5.1	Danksagungen	30
5.2	Änderungshistorie	30
5.3	lernOS Glossar	31

1 Über lernOS

Hinweis: Das ist die Baustelle auf der gerade der lernOS KI Leitfaden entsteht. Version 0.1. ist voraussichtlich Ende 2023 verfügbar. Wer sich für die Entstehungsgeschichte interessiert oder gar mitmachen will, findet alle Infos in diesem Beitrag auf CONNECT¹.

Ziel des Leitfadens: Menschen ohne KI-Vorbildung an das Thema Künstliche Intelligenz heranführen, damit sie eine informierte Entscheidung treffen können, wie/wo sie betroffen sind und welchen Nutzen sie durch KI haben könnten. Die Technologien/Beispiele sollen in und außerhalb von Organisationen funktionieren (Internet und Intranet).

Zielgruppe: Anwender:innen (Menschen, die vor der “Mattscheibe” sitzen) nicht Entwickler:innen; Anwender sollten aber Hintergründe verstehen. Kontext: Menschen in und außerhalb von Organisationen?

Ergänzung: Am 08. September hatten wir die Formulierung gefunden: Menschen vor der Mattscheibe mit der Offenheit und dem Interesse dahinterzublicken.



lernOS ist eine Methode zur Selbstorganisation für Menschen, die im 21. Jahrhundert leben und arbeiten. Um heute erfolgreich zu sein, muss man ständig lernen, sich organisieren und weiterentwickeln. Niemand sonst ist für diesen Prozess verantwortlich. Man muss sich selber darum kümmern (selbstgesteuertes, lebenslanges Lernen).

¹<https://community.cogneon.de/t/ein-lernos-leitfaden-zu-kuenstlicher-intelligenz-in-der-praxis-von-wissensbeiterinnen/3872>

lernOS Leitfäden stehen unter der Lizenz Creative Commons Namensnennung 4.0 International² (CC BY 4.0):



Du darfst:

- **Teilen** - das Material in jedwedem Format oder Medium vervielfältigen und weiterverbreiten.
- **Bearbeiten** - das Material remixen, verändern und darauf aufbauen und zwar für beliebige Zwecke, sogar kommerziell.

Unter folgenden Bedingungen:

- **Namensnennung** - Du musst angemessene Urheber- und Rechteangaben machen, einen Link zur Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden. Diese Angaben dürfen in jeder angemessenen Art und Weise gemacht werden, allerdings nicht so, dass der Eindruck entsteht, der Lizenzgeber unterstütze gerade Sie oder Ihre Nutzung besonders.
- **Keine weiteren Einschränkungen** - Du darfst keine zusätzlichen Klauseln oder technische Verfahren einsetzen, die anderen rechtlich irgendetwas untersagen, was die Lizenz erlaubt.

2 Grundlagen

Test

2.1 Grundidee der lernOS Leitfäden

Ein lernOS Leitfaden hilft Lernenden durch einen vorgegebenen Lernpfad ihr eigenes Lernziel im Rahmen eines Learning Sprints (3 Monate) zu erlernen zu erreichen. Die Lernenden lernen dabei allein, im Lerntandem oder in einem Learning Circle (4-5 Personen). Der zeitliche Aufwand über den Sprint sollte maximal zwei Stunden pro Woche betragen. Das Lernziel kann dabei aus einem oder einer Kombination dieser Bereiche kommen:

- **MINDSET** - eine bestimmte Haltung entwickeln
- **SKILLSET** - eine Fähigkeit erlernen
- **TOOLSET** - ein Tool oder eine Methode beherrschen

Die Woche 0 und 12 im Learning Sprint sind immer für Planung und Retrospektive vorgesehen, in Woche 4 und 8 findet ein sog. Boxenstopp zur Zwischenreflexion des Lernfortschritts statt. Die Wochen

²<https://creativecommons.org/licenses/by/4.0/deed.de>

1-11 können mit Übungen (mindestens 11), Katas³ genannt, gefüllt werden. Eine Kata sollte einen Übenden zw. 30-60 Minuten Zeit kosten und auf nicht mehr als zwei DIN A4 Seiten beschrieben sein (Sushi-Card-Prinzip).

Ein lernOS Leitfaden steht immer unter der Creative Commons Namensnennung 4.0 International Lizenz⁴ (CC BY). Die Lizenz gilt auch für im Leitfaden verwendete Bilder. Alle Inhalte müssen auch in bearbeitbaren Quellformaten vorliegen (bei Bildern z.B. PNG und SVG). Die Inhalte eines lernOS Leitfadens werden als Repository auf GitHub⁵ verwaltet, um offenen Zugang und einfach Wiederverwendung zu gewährleisten.

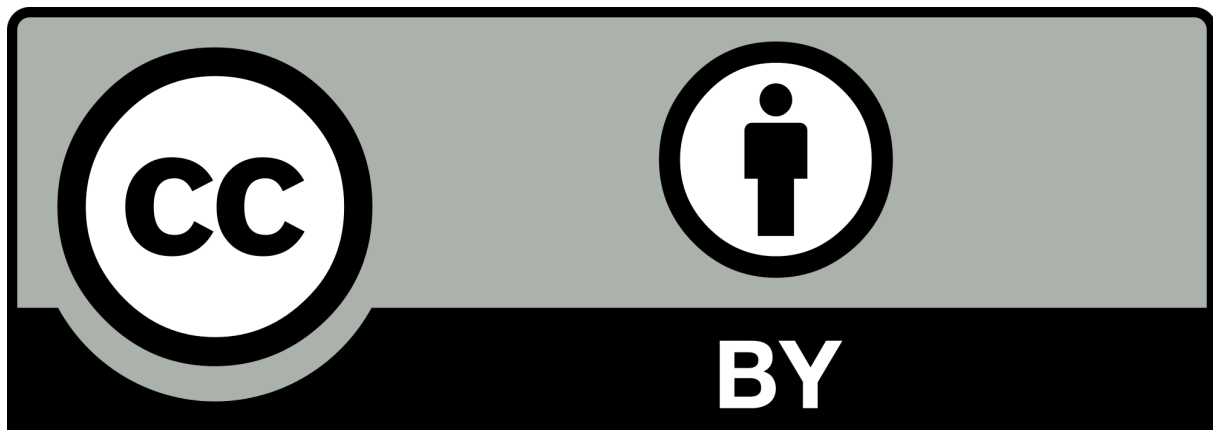


Abbildung 1: Offene Lizenz Creative Commons Namensnennung 4.0 International

Dadurch können lernOS Inhalte für jeden Zweck offen zugegriffen, genutzt, verändert und geteilt werden (auch für kommerzielle Anwendungen). Produkte und Dienstleistungen dürfen allerdings nicht den Begriff “lernOS” im Namen tragen. Dieser Ansatz ist vergleichbar mit dem Browser Chromium⁶ und den darauf aufsetzenden Produkten Google Chrome⁷, Microsoft Edge⁸, Opera⁹ und Brave.

2.2 Geschichte der Künstlichen Intelligenz

Die Geschichte der Künstlichen Intelligenz reicht zurück bis in die 1950er Jahre. Die folgende Tabelle gibt dir einen Überblick über die wichtigsten Meilensteine:

³[https://de.wikipedia.org/wiki/Kata_\(Programmierung\)](https://de.wikipedia.org/wiki/Kata_(Programmierung))

⁴<https://creativecommons.org/licenses/by/4.0/deed.de>

⁵<https://github.com>

⁶[https://de.wikipedia.org/wiki/Chromium_\(Browser\)](https://de.wikipedia.org/wiki/Chromium_(Browser))

⁷https://de.wikipedia.org/wiki/Google_Chrome

⁸https://de.wikipedia.org/wiki/Microsoft_Edge

⁹[https://de.wikipedia.org/wiki/Opera_\(Browser\)](https://de.wikipedia.org/wiki/Opera_(Browser))

Jahr Meilenstein

- 1956** Der Dartmouth Workshop¹⁰ ist die Geburtsstunde der Künstlichen Intelligenz als Fachgebiet.
- 1959** Allen Newell und Herbert A. Simon entwickeln den Logic Theorist¹¹, das erste KI-Programm.
- 1967** Dendral¹² wird entwickelt, ein regelbasiertes System für die chemische Analyse, eine bedeutende KI-Leistung.
- 1969** Shakey the Robot¹³ wird der erste mobile Roboter, der logisch denken und Probleme lösen kann.
- 1973** Der KI-Winter¹⁴ beginnt aufgrund hoher Erwartungen und unerfüllter Ziele in der KI-Forschung.
- 1980er** Expertensysteme¹⁵ gewinnen an Popularität. Sie verwenden Regeln, um menschliches Fachwissen in engen Bereichen zu imitieren.
- Jahre**
- 1997** IBM Deep Blue¹⁶ besiegt den Schachweltmeister Garri Kasparow und demonstriert damit das Potenzial der KI.
- 2011** IBM Watson¹⁷ gewinnt die Spielshow Jeopardy! und demonstriert damit die natürliche Sprachverarbeitung der KI.
- 2012** Geoffrey Hinton's Deep Learning¹⁸-Techniken lassen das Interesse an neuronalen Netzwerken wieder aufleben.
- 2014** Google DeepMind¹⁹ entwickelt ein neuronales Netzwerk, das lernt, Videospiele zu spielen.
- 2016** AlphaGo²⁰ von DeepMind besiegt den Go-Weltmeister Lee Sedol und beweist damit das strategische Denken der KI.
- 2017** Generative Adversarial Networks²¹ (GAN) erlangen Aufmerksamkeit für die Erstellung naturgetreuer Bilder.

¹⁰https://en.wikipedia.org/wiki/Dartmouth_workshop

¹¹https://en.wikipedia.org/wiki/Logic_Theorist

¹²<https://en.wikipedia.org/wiki/Dendral>

¹³https://en.wikipedia.org/wiki/Shakey_the_robot

¹⁴https://en.wikipedia.org/wiki/AI_winter

¹⁵https://en.wikipedia.org/wiki/Expert_system

¹⁶[https://en.wikipedia.org/wiki/Deep_Blue_\(chess_computer\)](https://en.wikipedia.org/wiki/Deep_Blue_(chess_computer))

¹⁷https://en.wikipedia.org/wiki/IBM_Watson

¹⁸https://en.wikipedia.org/wiki/Deep_learning

¹⁹https://en.wikipedia.org/wiki/Google_DeepMind

²⁰<https://en.wikipedia.org/wiki/AlphaGo>

²¹https://en.wikipedia.org/wiki/Generative_adversarial_network

Jahr Meilenstein

2021 Ameca²² ist ein humanoider Roboter, der von Engineered Arts entwickelt wurde. Ameca ist in erster Linie als Plattform für die Weiterentwicklung von Robotik-Technologien für die Mensch-Roboter-Interaktion gedacht. Die Interaktion kann entweder durch GPT-3 oder menschliche Telepräsenz gesteuert werden.

2021 Die generative KI DALL-E²³ zur Generierung von Bildern aus Text wird veröffentlicht.

2022 Der Chatbot ChatGPT²⁴, der das Large Language Model GPT verwendet, wird veröffentlicht.

2.2.1 ToDo

- Künstliche Intelligenz²⁵
Geschichte aus OpenHPI MOOC ergänzen

2.3 ## Künstliche Intelligenz und Maschinelles Lernen

Stellen wir uns vor, ein Kind sieht zum ersten Mal einen Hund. Das Kind ist fasziniert von diesem neuen Tier und ruft begeistert "Wau Wau!". In seiner kindlichen Begeisterung bezeichnet es zunächst jeden Vierbeiner als "Wau Wau". Erst nach und nach lernt es durch Beobachtung und mit Hilfe seiner Eltern, die feinen Unterschiede zwischen verschiedenen Tieren wie Hunden und Katzen zu erkennen.

Dieser Prozess der Erkundung und Anpassung findet sich auch beim maschinellen Lernen wieder. Computerprogramme lernen aus Beispielen, ohne explizit programmiert zu sein. Am Anfang kennen sie die feinen Unterschiede zwischen verschiedenen Kategorien noch nicht. Durch das Analysieren großer Mengen von Daten (das "Training") lernen sie nach und nach, Muster zu erkennen und Daten korrekt einzuordnen.

Der gesamte Prozess kann als Kreislauf von insgesamt 7 Schritten dargestellt werden. Wichtig ist, dass der Lernprozess nach einem Durchlauf nicht abgeschlossen ist, sondern in mehreren Schleifen erfolgt. Auch das ist eine Gemeinsamkeit zwischen dem menschlichen und dem maschinellen Lernen.

Nehmen wir zur Erläuterung des gesamten Lernprozesses ein konkretes Beispiel: Eine KI-Anwendung soll lernen, Bilder von Hunden und Katzen zu unterscheiden.

²²[https://en.wikipedia.org/wiki/Ameca_\(robot\)](https://en.wikipedia.org/wiki/Ameca_(robot))

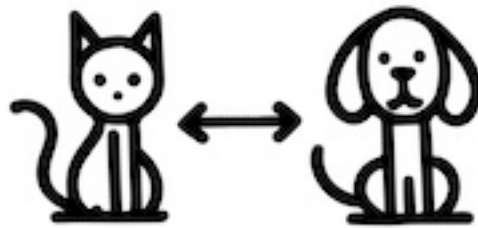
²³<https://en.wikipedia.org/wiki/DALL-E>

²⁴<https://en.wikipedia.org/wiki/ChatGPT>

²⁵https://de.wikipedia.org/wiki/K%C3%BCnstliche_Intelligenz



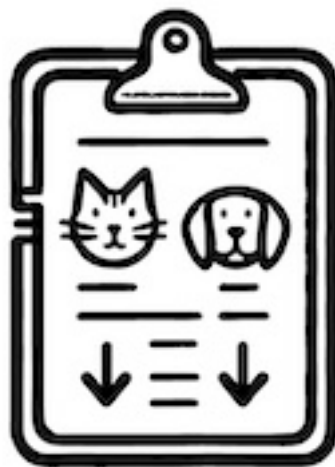
Abbildung 2: Untitled



1. Problemdefinition

Bei Kindern ist die “Problemdefinition” nicht so explizit wie in einem KI-Projekt, sondern eher durch eine intrinsische Motivation ausgelöst oder durch ein von der Umwelt stimulierter Lernanreiz. Kinder sind neugierig und wollen die Welt um sich herum verstehen. Im unserem konkreten Beispiel möchte ein Kind lernen, die Unterschiede zwischen einem Hund und einer Katze zu erkennen.

Generalisiert übertragen auf die KI heißt das, dass sie Bilder klassifizieren soll.

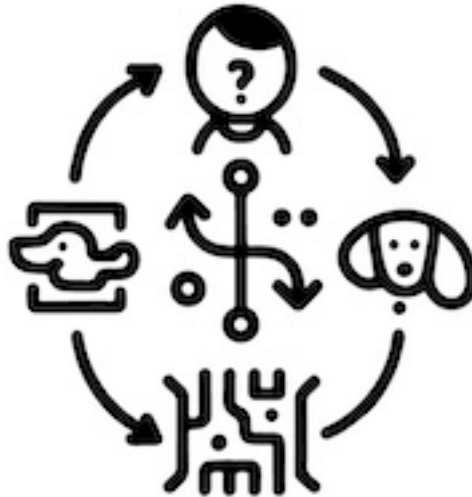


2. Datensammlung

In der realen Welt sieht das Kind viele verschiedene Hunde und Katzen. Es erkennt Tiere, wenn sie 4 Beine haben und einen Schwanz und sagt “Wau wau”. Die Eltern helfen dem Kind, indem sie auf die Tiere zeigen und diese benennen. So verbindet das Kind die Bilder und Laute der Tiere mit den richtigen Begriffen.

Für die KI werden Tausende von Bildern von Hunden und Katzen gesammelt und entsprechend gekennzeichnet. Diese Daten müssen eventuell bereinigt werden, um Fehler oder irrelevante

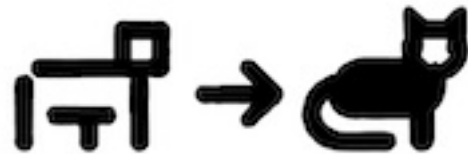
Informationen zu entfernen.



3. Modellauswahl

Im Kopf des Kindes bilden sich durch Wiederholung und Korrektur feste neuronale Verbindungen, was zu einer klareren Unterscheidung zwischen Hunden und Katzen führt. Dieser Prozess ist ähnlich der Art und Weise, wie ein **neuronales Netz** (vgl. das Kapitel...in diesem Leitfaden) im KI-Training gestärkt wird. Es ist besonders geschickt darin, in unstrukturierten Daten Muster zu erkennen und aus diesen zu lernen.

Mit jedem durchlaufenen Trainingsbild verbessert das KI-Modell seine Erkennungsfähigkeit für charakteristische Merkmale wie die Beschaffenheit des Fells, die Form der Ohren und die Beschaffenheit des Schwanzes. Es optimiert seine Vorhersagegenauigkeit, indem es relevante Muster hervorhebt und weniger wichtige vernachlässigt. Diese schrittweise Verfeinerung seiner Erkennungsleistung gleicht dem Lernprozess eines Kindes, das durch ständiges Ausprobieren und die daraus resultierenden Korrekturen lernt.



4. Training

Das Kind lernt durch Wiederholung und Rückmeldung der Eltern. Wenn es einen Hund als Katze bezeichnet, wird es korrigiert. So verfeinert das Kind mit jeder Korrektur sein inneres Modell.

In der KI gibt es zwei grundlegende Arten des Trainings:

- Beim überwachten Lernen (**Supervised Learning**) bekommt das Modell zu allen Trainingsdaten die korrekte Klassifikation. Anhand dieser Labels lernt es die relevanten Merkmale. In diese Lernkategorie gehören auch neuronale Netze.
- Beim unüberwachten Lernen (**Unsupervised Learning**) erhält das Modell nur die Daten ohne Labels. Es muss selbst Ähnlichkeiten erkennen und die Daten gruppieren. So entdeckt es eigenständig Muster und Strukturen. Das unüberwachte Lernen kann als ergänzende Methode beim Training eingesetzt werden, um dem Modell noch tiefergehendes Verständnis zu ermöglichen. Es hilft dem Modell, latente Merkmale und Zusammenhänge zwischen den Daten zu erfassen, die möglicherweise nicht durch das überwachte Lernen allein erfasst werden könnten.

Die Ergänzung bzw. die Kombination von neuronalen Netzen mit dem überwachten Lernen nennt man **Deep Learning**, ein Begriff, der auch häufig in der KI-Diskussion auftaucht.

So wie ein Kind durch Wiederholung und Korrektur seine Unterscheidungsfähigkeit verbessert, optimiert das KI-Modell seine Leistung iterativ durch viele Trainingsdurchläufe und Anpassungen. Nach genügend Training kann es dann auch neue Daten zuverlässig klassifizieren.



5. Validierung

Ähnlich wie ein Kind lernen muss, Hunde und Katzen auch in neuen Situationen richtig zu erkennen, muss ein KI-Modell beweisen, dass es in der Lage ist, Daten zu generalisieren und nicht nur auswendig gelernt hat. Dazu werden neue Testdaten verwendet, um die Fähigkeit des Modells zur korrekten Klassifikation zu überprüfen. Wenn die Ergebnisse unzureichend sind, muss das Modell verbessert werden, um die zugrunde liegenden Regeln zu verstehen und nicht nur einzelne Merkmale auswendig zu lernen. So wie Eltern ein Kind herausfordern und bei Bedarf korrigieren, hilft die Validierung dabei, das KI-Modell zu testen und zu verbessern. Dieses Konzept ist entscheidend beim maschinellen Lernen und ermöglicht es dem Modell, seine Leistung kontinuierlich zu verbessern.



6. Anwendung

Nach dem Training kann das KI-Modell sein gelerntes Wissen anwenden, ähnlich wie ein Kind sein Wissen nutzt, um neue Tiere draußen korrekt zu erkennen und zu benennen. Das trainierte Modell kann in verschiedenen Anwendungen eingesetzt werden, wie z.B. einer Bildanalyse-App oder einem Empfehlungssystem. Es wendet sein gelerntes Wissen auf neue Daten an und ermöglicht es, nützliche Aufgaben zu lösen. Genau wie ein Kind verschiedene Tiere erkennt und dieses Wissen praktisch anwendet, findet ein trainiertes KI-Modell Anwendung in realen Systemen.



7. Feedback

Regelmäßiges Feedback ist für das Kind unerlässlich, um zu lernen und sich weiterzuentwickeln. So wird es beispielsweise durch Spiele, die die Unterscheidung von Farben und Formen fördern, oder durch interaktive Bildungstools, die spielerisch Wissen vermitteln, weiter gefordert und unterstützt. Gleiches gilt für ein KI-Modell, das durch die kontinuierliche Analyse von Nutzerinteraktionen in einer Foto-Sortier-App oder durch die Einbeziehung neuer, vielfältiger Bilddatensätze stetig verfeinert wird. So wie ein Kind durch den Besuch eines Zoos oder durch das Blättern in einem Tierbuch neue Tierarten kennenlernt und sein Wissen vertieft, erweitert das KI-Modell seine Erkennungsfähigkeiten durch die Einführung zusätzlicher, unterschiedlicher Bilder oder durch Feedback von Benutzern, die Fehlklassifikationen melden. Diese fortlaufende Interaktion, sei es durch menschliche Rückmeldung oder durch neue Dateninputs, ermöglicht es dem Modell, adaptiv zu bleiben und sich an die sich wandelnde Welt anzupassen.

Reflexionsfragen:

Welche Meinung hast du selbst zu folgenden Punkten, die im Zusammenhang mit der weiteren Entwicklung des maschinellen Lernens diskutiert werden? Welche Zukunftsaspekte stimmen Dich skeptisch, welchen stehst du eher positiv gegenüber, wenn sie so eintreffen? Glaubst du an diese Potentiale?

1. **Lernprozesse und Anpassungsfähigkeit:** Maschinelle Lernmodelle werden in der Lage sein,

nicht nur die Komplexität menschlicher Lernprozesse nachzubilden, sondern sich auch an neue Situationen anzupassen, indem sie emotionale und soziale Kontexte integrieren und flexibel auf Veränderungen reagieren.

2. **Generalisierung und Transferlernen:** Fortschritte im maschinellen Lernen ermöglichen es Systemen, mit minimalen Datenmengen zu generalisieren und Wissen über verschiedene Domänen hinweg zu transferieren, ähnlich der menschlichen Kapazität, aus wenigen Beispielen zu lernen und Erkenntnisse in unterschiedlichen Kontexten anzuwenden.
3. **Autonome Motivation und kontextuelles Verständnis:** Zukünftige maschinelle Lernmodelle werden ein tiefes Verständnis für Kontext und Nuancen erlangen und eine eigene Form der "Motivation" entwickeln, was sie in die Lage versetzt, kontextbewusst und autonom zu handeln.
4. **Interaktives Lernen und kontinuierliche Verbesserung:** Durch die Integration von interaktiven und sozialen Feedbackmechanismen werden maschinelle Lernsysteme eine Lernerfahrung bieten, die kontinuierliche Verbesserung ermöglicht und der menschlichen Lernerfahrung ähnelt.

2.4 Neuronale Netzwerke

- Input/Output Layer, Hidden Layers
- Lineare Algebra
- Training (Gewichte, Bias, RLHF, HIL) -> Aspekte Nachhaltigkeit/Ethik
- Refinement
- Transformer (Tokens, Kontext, Decoder, Encoder, Embedding)
- Quantization/Kompression

Videos:

- Aber was *ist* nun ein neuronales Netzwerk? | Teil 1, Deep Learning²⁶
- Gradient descent, how neural networks learn | Chapter 2, Deep learning²⁷

2.5 KI-Anwendungsfelder

- Aus huggingface/models²⁸ ein vereinfachtes Modell ableiten
- Brücke zu Aufgaben von Wissensarbeiter:innen bauen (BIBB, WINS)
- Chatbots

Simon: fände eine Infografik in diesem Stil cool.

²⁶<https://www.youtube.com/watch?v=aircAruvnKk>

²⁷<https://www.youtube.com/watch?v=IHZwWFHwa-w>

²⁸<https://huggingface.co/models>

2.6 KI Modelle

Das Kapitel gibt einen Überblick über KI-Modelle und unterteilt sie in verschiedene generative Lösungen und Anwendungen. Dies lässt besser verstehen, wie diese Modelle in bestimmten Szenarien eingesetzt werden. Es ist eine Hilfe, wenn für eigene Projekte oder geschäftliche Anforderungen das am besten geeignete Modell zu wählen ist.

2.6.1 Wieso sprechen wir von Modellen?

Für das Ergebnis von Anwendungen, die mit generativer KI arbeiten, ist das zugrunde liegende so genannte Modell wichtig. Es ist ein Modell, eine Abstraktion, ein aus Trainingsdaten “gelerntes” Regelwerk von Strukturen, Regelmäßigkeiten und Wahrscheinlichkeiten. Eine Sprache beispielsweise umfassend formal zu erlernen, ist für eine zeitnahe Verarbeitung in Dialogen (noch) deutlich zu komplex. So ist zu verstehen, wie stark die Interpretationen der Eingaben von Nutzer:innen und die darauf angewandte Entwicklung von Antworten oder Ergebnissen von den Modellen und diese wiederum auf den ausgewählten Daten in der Trainingsphase abhängen.

Warum ist es wichtig, das Modell der KI zu kennen? Nun, es ist ein Modell. Ein Modell stellt immer eine Vereinfachung, eine Näherung dar.

“ChatGPT Is a Blurry JPEG of the Web”

Science Fiction-Autor Ted Chiang in “The New Yorker”

Ein guter Vergleich, wenn wir an den Qualitätsverlust von Grafik-, Audio- oder Video-Daten durch Kompressionsformate wie JPG, MP3 oder MPEG denken. Die Abstraktion der Originaldaten wird durch Unschärfen oder blechnen Klang spürbar. Dieser Vergleich ist ein wichtiger Merkposten, wenn eine generative KI-Anwendung grammatikalisch und stilistisch perfekt klingende Texte kreiert. Es bleibt eine unscharfe Annäherung, wir beobachten so genannte “Halluzinationen” mit inhaltlich unsinnigen Antworten auf sachliche Fragen.

Nehmen wir als Beispiel KI-Anwendungen, die Texte verstehen und generieren. Die technische Grundlage dieser Lösungen sind Sprachmodelle, die ein lernendes System auf der Basis angebotener Texte in sehr großer Menge entwickelt hat. Dazu wurde mit Software-Algorithmen aus dem Gebiet des Natural Language Processing gearbeitet, denn diese sind geübt in statistischen Vorhersagen von möglichen nächsten Worten für einen stimmigen Text. Ihnen wurde eine Sprache nicht “beigebracht”, sondern sie haben die Inhalte entsprechend ausgewertet, um Semantik, Grammatik, Regeln und Regelmäßigkeiten einer Sprache zu erkennen.

Ein Beispiel zur Erkennung von Objekten: Wird eine KI zur Erkennung von Katzen in Fotos trainiert, werden dazu viele Bilder von Katzen bereitgestellt, um sie selbstlernend auf Regelmäßigkeiten und

Regeln zu analysieren. Sind das nun Katzenfotos, die immer die gleiche gelbe Augenfarbe haben, wird das sehr wahrscheinlich als eine Regel zur Unterscheidung von anderen Tierfotos gespeichert. Zeigen wir der KI dann in der Praxis das Foto einer Katze mit blauen Augen, wird sie diese nicht erkennen.

2.6.2 Überblick zu Modellen generativer KI

Diese Übersicht zeigt beispielhaft einige KI-Modelle und ihre Anwendungen. Derzeit sind die Instruktionen und Bewertungsmodelle, die in den meisten KI-Modellen verwendet werden, urheberrechtlich geschützt und nicht für die Öffentlichkeit einsehbar. Auch entwickeln und trainieren die verschiedenen Anbietenden ihre KIs weiter. Daher ist die Versionsbezeichnung der jeweiligen Modelle essentiell für die qualitative Einschätzung der generierten Inhalte. Insbesondere wenn der Umfang der Trainingsdaten erheblich vergrößert wird.

Bei der Einteilung von KI-Modellen auf der Grundlage ihrer generativen Verwendungszwecke können wir die Texterstellung inklusive Code-Generierung, die Bilderstellung und die Audio-/Videoerstellung als unterschiedliche Kategorien betrachten.

Modelle zur Textgenerierung / Code

KI-Modelle können eingegebene Texte (oder gesprochene Sprache) verstehen und daraufhin neue Texte generieren, die einem von Menschen geschriebenen Inhalt ähneln. Das können Sprachübersetzungen, Textverbesserungen, Chatbot-Dialoge, Stilübertragungen oder die Generierung von Inhalten wie Gliederungen, Blogbeiträgen, Artikeln, Kursfragen sein. Genauso schreiben entsprechend trainierte KIs auch Software-Code.

- **Luminous**²⁹ von Aleph Alpha ist ein in fünf europäischen Sprachen trainiertes Sprachmodell: Deutsch, Englisch, Französisch, Italienisch und Spanisch. Als Eingabe zur Textentwicklung kann mit Text oder kombiniert mit Bildern erfolgen.
- **GPT-4**³⁰ (Generative Pre-trained Transformer) von OpenAI ist das bekannteste Modell für Dialoge, Textgenerierung oder auch Code-Entwicklung. Die Eingabe kann mit Text oder Bildmaterial erfolgen. Das Sprachmodell wird iterativ durch Verstärkungslernen mit menschlichem Feedback (Reinforcement Learning from Human Feedback, RLHF) verbessert.
- **BARD**³¹ ist Googles experimentelles Sprachmodell mit den Varianten *LaMDA* (Language Models for Dialog Applications; Fokus auf Dialoge) und *PaLM 2* (Pathways Language Model; spezialisiert auf anspruchsvolle logische Aufgaben einschließlich Programmierung und Mathematik, Klassifizierung und Beantwortung von Fragen, Übersetzung und Mehrsprachigkeit sowie zur Generierung natürlicher Sprache).

²⁹<https://docs.aleph-alpha.com/docs/introduction/luminous/>

³⁰<https://openai.com/gpt-4>

³¹<https://bard.google.com/chat>

- **Llama 2**³² von Meta mit den Modellvarianten *Llama Chat*, ein Sprachmodell, das iterativ durch Verstärkungslernen mit menschlichem Feedback (RLHF) für Dialoge verfeinert wird, und *Code Llama*, ein Sprachmodell, das mit code-spezifischen Datensätzen für die Softwareentwicklung trainiert wird, um Code zu schreiben, zu vervollständigen oder das Debugging zu unterstützen.
- **LEAM**³³ (Large European AI Models) in Entwicklung befindliche KI-Modelle, die europäische Werte und hohe Anforderungen in Bezug auf Datenschutz, Transparenz und Voreingenommenheit widerspiegeln. Gleichzeitig soll die Gestaltung des Trainings von KI-Modellen nachhaltiger werden.
- **LeoLM**³⁴ (Linguistically Enhanced Open Language Model), erste umfassende Suite deutschsprachiger Sprachmodelle.
- **Whisper**³⁵: Universell einsetzbares Modell von Open AI, das die mehrsprachige Erkennung von Sprache in Audiodateien bietet und das Ergebnis als Text oder übersetzten Text ausgibt.

Modelle zur Bilderzeugung

KI-Modelle können auf der Basis von Texteingaben (zum Teil auch Bilddateien) neue Bilder erzeugen, die realen Objekten oder Szenen ähneln. Dazu gehören Aufgaben wie Bildsynthese, Stilübertragung, oder Bildverbesserung (Superresolution).

Sie nutzen die so genannte Diffusion, was die Verteilung von Partikeln im Raum beschreibt. Diesem ähnlich verändert die KI einzelne Pixel in einem Bild fortlaufend und in Interaktion miteinander basierend auf gelernten Informationen, um neue Inhalte zu erzeugen.

- **DALL•E**³⁶ von Open AI versteht Beschreibungen in natürlicher Sprache, um daraufhin real wirkende Fotos und Kunstwerke zu erstellen. Die Version *DALL-E 2* bietet die Option, ein bestehendes Bild zu bearbeiten oder Variationen eines bereitgestellten Bildes zu erstellen. *DALL•E 2 und 3* bilden in verschiedenen Anwendungen die KI-Basis.
- **Firefly Image 1**³⁷ von Adobe setzt nach eigenen Angaben auf lizenzierte Fotos der eigenen Bilddatenbank und gemeinfreies Bildmaterial. Für die Version 2 ist das individuelle Trainieren mit eigenen Werken in der Entwicklung.
- **LoRA Stable Diffusion** (Learn On Reconstruction and Attention) ist eine Kombination von Algorithmen für das Fine-Tuning von Bildern und Bildstil-Training. Nach Training mit ausgewählten Bildern erkennt die KI einen bestimmten Stil, um diesen dann auf andere Bilddaten anzuwenden.
- **Stable Diffusion**³⁸ ist eine der wenigen KI-Anwendungen, die Methoden und Trainingsdaten

³²<https://ai.meta.com/llama/>

³³<https://leam.ai/>

³⁴<https://laion.ai/blog-de/leo-lm/>

³⁵<https://platform.openai.com/docs/models/whisper>

³⁶<https://platform.openai.com/docs/models/dall-e>

³⁷<https://firefly.adobe.com>

³⁸<https://stablediffusionweb.com/>

öffentlich frei zugänglich machen. Die Bildgenerator-KI nutzt *LAION-5B*³⁹, der mit fast 6 Milliarden Verweisen auf Bilder mit Beschreibung im Internet (LAION - Large-scale Artificial Intelligence Open Network, 5B - 5 billion text-image pairs) größte öffentlich zugängliche Trainingsdatensatz (Du kannst hier suchen, ob deine Fotos ebenfalls für Trainings verwendet wurden: Have I Been Trained?⁴⁰)

Modelle zu Audio-/Videogenerierung

Solche KI-Modelle machen es Anwendungen möglich, natürlich wirkende Sprache zu synthetisieren, Musik zu komponieren, Videos oder Deepfakes zu generieren. Künstlich erzeugte Sprachausgaben gibt es schon lange, sie klingen selten wie eine Sprechstimme. Nun wird die Spachsynthese durch KI auf ein neues Niveau gebracht:

- **TTS**⁴¹: Text-to-Speech-KI-Modell von Open AI, das nach Texteingabe natürlich klingende Sprache ausgibt. Die Variante *tts-1* wurde für Anwendungsfälle in Echtzeit optimiert. Für *tts-1-hd* wird der Fokus auf Qualität gelegt.
- **English v1**⁴²: Englischsprachiges Sprachsynthese-Modell von ElevenLabs.
- **Multilingual v2**⁴³: Sprachsynthese-Modell von ElevenLabs für 28 Sprachen.

2.6.3 Hinweis

Bitte habe diese Diskussionen zu Modellen und ihren Trainingsdaten im Blick:

- Es wurden Rechtsklagen eingereicht von Kunstschaffenden gegen Stability AI oder Midjourney, weil diese urheberrechtlich geschützte Werke im Training genutzt haben.
- Softwareentwickler klagen gegen Unternehmen wie GitHub, Microsoft oder OpenAI, weil diese ihren Open-Source-Code als Trainingsdaten für die KI-Entwicklung hernehmen.

2.7 KI-Tools und -Dienste

In diesem Kapitel geben wir dir einen Überblick, wie du mit verschiedenen Tools und Diensten das KI-unterstützte Arbeiten kennenlernen kannst. Für die Auswahl haben wir uns jeweils folgende Fragen gestellt:

- Ist das Tool / der Dienst für den Einstieg nutzbar?
- Gibt es eine kostenfreie Nutzung?

³⁹<https://laion.ai/>

⁴⁰<https://haveibeenentrained.com/>

⁴¹<https://platform.openai.com/docs/models/tts/>

⁴²<https://elevenlabs.io/docs/speech-synthesis/models/>

⁴³<https://elevenlabs.io/docs/speech-synthesis/models/>

- Sind die Benutzungsoberfläche und erklärende Hilfestellungen auch in deutscher Sprache verfügbar?
- Können wir davon ausgehen, dass die Vorgaben der DSGVO berücksichtigt werden?
- Welche 3 bis 5 sind aktuell unsere besten Empfehlungen?

(Oft ist zumindest eine Registrierung erforderlich.)

2.7.1 Tipps & Empfehlungen

Hier ist so viel in Bewegung, das Kapitel könnten wir täglich ergänzen oder erweitern. Kostenfreie Angebote setzen unter Umständen ältere KI-Anwendungen ein, die weniger leistungsfähig oder aktuell im Informationsstand sind. Wer kostenpflichtige Tools nutzt, hat zudem oft einen größeren Funktionsumfang.

Bei manchen Tools und Diensten ist es trotz deutschsprachiger Oberfläche sinnvoll, die Eingaben in englischer Sprache zu machen. Die KIs sind oft mit englischsprachigen Daten trainiert worden.

Bekannte Office-Anwendungen wie Microsoft 365 bieten der Unternehmens-Kundschaft nach und nach entsprechende Erweiterungen oder Integrationen an. Habe diese Anbieter auf jeden Fall im Blick:

- **ChatGPT**⁴⁴: Kommunikation via Text, Audio und Bildeingabe sowie Bildgenerierung in kostenpflichtiger Plus-Version.
- **Microsoft 365 Copilot**⁴⁵: KI-gestützte Assistenz in den Microsoft Produkten für Texterstellung, Zusammenfassungen, Textüberarbeitungen, Bildersuche oder Generierung von Präsentationen.

2.7.2 KI-gestützte Multi-Tools

Diese Angebote bieten gleich mehrere Werkzeuge für die verschiedenen Anwendungsgebiete unter einem Dach.

- **Bing Chat / Copilot**⁴⁶: Kommunikation via Text, Audio und Bildeingabe sowie Code-Generierung auf der Basis von *GPT-4*; Bildgenerierung mit *Dall-E 3*; Funktionsumfang abhängig vom Betriebssystem, deutschsprachig, kostenfreie Basis-Nutzung.
- **Fobizz Tools**⁴⁷: Tools für Schule und Weiterbildung für Chats mit Persönlichkeiten der Geschichte, Textzusammenfassungen, gesprochene Inhalte in Text wandeln, Handschrifterkennung, Bildgenerierung, kostenfreie Basis-Nutzung, DSGVO-konform. Mit einer Anmeldung verschiedene

⁴⁴<https://chat.openai.com/auth/login>

⁴⁵<https://support.microsoft.com/de-de/copilot>

⁴⁶<https://www.bing.com/search>

⁴⁷<https://tools.fobizz.com>

AI-Modelle testen (Bild: *Dall-E 2, Stable Diffusion*; Text & Chat: *Luminous Extended, GPT3.5 turbo, GPT4, Open Assistant, Claude 2*).

- **Neuroflash**⁴⁸: *ChatFlash, ContentFlash, ImageFlash* und *PerformanceFlash* für Texterstellung, Textüberarbeitung, Bilderzeugung, deutschsprachig, kostenfreie Basis-Nutzung, DSGVO-konform.
- **Poe**⁴⁹: Mit einer Anmeldung verschiedene AI-Modelle im Chat testen (aktuell u.a. *ChatGPT & GPT-4, Claude Instant & Claude 2, StableDiffusionXL, PaLM, Llama 2*), kostenfreie Basis-Nutzung.
- **You.Com**⁵⁰: Tools *YouChat, YouCode, YouImagine, YouWrite* für Recherche, Texterstellung, Code-Assistenz und Bildgenerierung, englischsprachig, kostenfreie Basis-Nutzung.

2.7.3 KI-gestützte Text-Tools

KI-Text-Tools ermöglichen die Interaktion mit dem Tool wie in einer Unterhaltung mit einem anderen Menschen. Sie unterstützen die Online-Recherche und geben vielfältige Hilfestellung bei der Texterstellung, ebenso bei Übersetzungen und Überarbeitungen.

- **AudioPen**⁵¹: Generiert aus mehrsprachigem, selbst undeutlichem oder stichpunktartigem Sprachinput ein- oder multisprachigen Text, kann in kostenpflichtiger Version Texte im Stil überarbeiten, kostenfreie Basis-Nutzung.
- **Bard**⁵²: KI-Experiment von Google, mit dem Übersetzungen und Texte generiert werden können, deutschsprachig, kostenfreie Basis-Nutzung.
- **Bing Chat / Copilot**⁵³: Kommunikation via Text, Audio und Bildeingabe auf der Basis von _GPT-4; _Funktionsumfang abhängig vom Betriebssystem, deutschsprachig, kostenfreie Nutzung.
- **BlogMojo**⁵⁴: Texterstellung mit Stil des individuellen Blogs, deutschsprachig, kostenfreie Basis-Nutzung, DSGVO-konform.
- **ChatGPT**⁵⁵: Textchat zur Textentwicklung, Übersetzung, Textbearbeitung etc. mit Datenstand bis Januar 2023, kostenfreie Basis-Nutzung.
- **Claude 2**⁵⁶: Kann in Europa via **Fobizz Tools** und **Poe** (s.o.) genutzt werden.
- **Deepl**⁵⁷: Übersetzungen und DeeplWrite zur Textverbesserung mit verschiedenen Stilauswahlen, deutschsprachig, kostenfreie Basis-Nutzung, DSGVO-konform.

⁴⁸<https://neuroflash.com/de/free-content-generatoren/>

⁴⁹<https://poe.com/>

⁵⁰<https://you.com>

⁵¹<https://audiopen.ai>

⁵²<https://bard.google.com/chat>

⁵³<https://www.bing.com/search>

⁵⁴<https://blogmojo.ai>

⁵⁵<https://chat.openai.com/auth/login>

⁵⁶<https://claude.ai/>

⁵⁷<https://deepl.com>

- **Luminous**⁵⁸: Sentiment-Analysen und Textgenerierung eingeteilt nach verschiedenen Anwendungsszenarien, englischsprachige Oberfläche, kostenfreie Basis-Nutzung, DSGVO-konform.
- **Pi**⁵⁹: Textchat mit Sprachausgabe zu verschiedenen Themen als empathisches Gegenüber, englischsprachig, kostenfreie Nutzung.

2.7.4 KI-gestützte Codierungshilfen

Code-Tools mit KI sind innovative Programme, die Entwickler:innen helfen, effizienter zu arbeiten. Sie generieren kurze Code-Snippets, eröffnen Diskussionen über verschiedene Lösungsansätze und bieten Erklärungen zu Code-Fragmenten. Außerdem können sie zur Überarbeitung bestehenden Codes eingesetzt werden, um dessen Qualität zu verbessern.

- **Bing Chat / Copilot**⁶⁰: Code-Assistenz für verschiedene Programmiersprachen, deutschsprachig, kostenfreie Basis-Nutzung.
- **ChatGPT**⁶¹: Code-Assistenz für verschiedene Programmiersprachen mit Fokus auf *Python*, deutschsprachig, kostenfreie Basis-Nutzung.
- **Github Copilot**⁶²: Code-Assistenz, die Vorschläge und Optimierungen für Softwarecode liefert, kostenfreie Testnutzung.
- **YouCode**: Code-Assistenz (s.o.)

2.7.5 KI-gestützte Bild-Tools

Bild-Tools nutzen KI für die Erzeugung und Optimierung von Fotos und Grafiken. Textanweisungen werden verarbeitet, verstanden und entsprechend für die Generierung von passenden Bildern verwendet. Sie sind in der Lage, so genannte Deepfakes zu erzeugen: Fotos (oder Videos) in denen die Gesichtszüge einer Person auf die Aufnahme eines anderen Menschen übertragen werden.

- **Adobe Express**⁶³: Generierung von Bildern sowie Bearbeitung und Verbesserung von Bildern, deutschsprachig, kostenfreie Basis-Nutzung.
- **Bing Creator**⁶⁴: Bildgenerierung mit *DALL-E 3*, deutschsprachig, kostenfreie Basis-Nutzung.
- **CivitAI**⁶⁵: Bildgenerierung mit Vorlagenauswahl und Text-Prompt, englischsprachig, Registrierung erforderlich, kostenfreie Basis-Nutzung.
- **Midjourney** (kostenpflichtig) via **Poe** nutzen (s.o.).

⁵⁸<https://app.aleph-alpha.com/jumpstart/>

⁵⁹<https://pi.ai/talk>

⁶⁰<https://www.bing.com/search>

⁶¹<https://chat.openai.com/auth/login>

⁶²<https://github.com/features/copilot>

⁶³<https://www.adobe.com/de/express/>

⁶⁴<https://bing.com/create/>

⁶⁵<https://civitai.com/>

2.7.6 KI-gestützte Audio-Tools

Audio-Tools mit KI erlauben die Spracheingabe zur Steuerung von Anwendungen. Sie können aufgrund von Textinputs natürlich wirkende Sprache erzeugen. Einige Tools ersetzen die Tonspur einer Videodatei, zum Beispiel mit einer Übersetzung des gesprochenen Textes, für das so genannte Dubbing. Andere Tools beherrschen das Voice Cloning, die Übertragung der Besonderheiten einer gesprochenen Stimme auf eine synthetisierte Sprachausgabe.

- **AudioPen**⁶⁶: Generiert aus mehrsprachigem, selbst undeutlichem oder stichpunktartigem Sprachinput ein- oder mehrsprachigen Text, kostenfreie Basis-Nutzung.
- **Descript**⁶⁷: Tool, das aus Sprachaufzeichnungen Zusammenfassungen und Social Media-Posts erzeugt, bietet mit *AI Voices* aus Texten erzeugte Sprache und Voice Cloning, kostenfreie Basis-Nutzung.
- **ElevenLabs**⁶⁸: Generiert aus Textinput hochwertige Audioqualität mit Emotionen in vielen Sprachen, große Auswahl an Stimmen, Dubbing von Videos, Voice Cloning und Hörbuch-Entwicklung, kostenfreie Basis-Nutzung.
- **Fliki**⁶⁹: Hochwertige deutschsprachige Stimmen, kostenfreie Basis-Nutzung. Voice Cloning in Premium-Tarif enthalten.
- **Murf**⁷⁰: Hochwertige deutschsprachige Stimmen, kostenfreie Basis-Nutzung, DSGVO-konform.

2.7.7 KI-gestützte Video-Tools

Video-Tools generieren mit KI aus Texteingaben passende Videos oder animieren menschenähnliche Avatare, deren Bewegung und Mimik sehr realistisch wirken. Durch die Übertragung von Besonderheiten eines realen Menschen auf eine Person in einem Video können einige Tools so genannte Deepfakes erzeugen, die enorm realitätsgetreu wirken.

- **Heygen**⁷¹: Video-Generierung mit menschlichen Avataren in vielen Sprachen aus Textskript, kostenfreies Demo-Video, DSGVO-konform.
- **Lumen5**⁷²: Erzeugen von Skripts und Videos auf Basis von Blogbeiträgen, Voice Cloning in beta, kostenfreie Basis-Nutzung.
- **Synthesia**⁷³: Mehrsprachige Generierung von Videos mit real wirkenden Avataren aus Texteingaben, kostenfreies Demo-Video.

⁶⁶<https://audiopen.ai>

⁶⁷<https://www.descript.com/>

⁶⁸<https://elevenlabs.io/>

⁶⁹<https://fliki.ai/>

⁷⁰<https://murf.ai/>

⁷¹<https://www.heygen.com/>

⁷²<https://lumen5.com/>

⁷³<https://www.synthesia.io/>

- **Pictory**⁷⁴: Wandelt Texteingaben in Videos um oder erstellt Kurzvideos aus Videoinhalten, kostenfreie Basis-Nutzung.

Viel Spaß beim Ausprobieren!

2.8 ## KI und Gesellschaft: eine Reflexion über Implikationen und Verantwortung

In der heutigen digitalen Ära stellt die Künstliche Intelligenz (KI) eine zentrale Technologie dar, die unsere sozialen und beruflichen Interaktionen neu gestaltet. Dieses Kapitel regt zur Reflexion und Diskussion über die Auswirkungen der KI im Kontext Deiner Organisation und der Gesellschaft an. Wir haben 7 Aspekte identifiziert, die im Zusammenhang mit „KI und Gesellschaft“ diskutiert werden. Zu jedem Aspekt sind Fragen zusammengestellt zur Selbstreflexion und zur Diskussion mit Deiner Lerngruppe.

Arbeitsplätze und Automatisierung

KI und Automatisierung werden viele Arbeitsbereiche und Berufsfelder grundlegend verändern. Dies löst sowohl Hoffnungen als auch Ängste aus. Diskutiert wird der Verlust von Arbeitsplätzen durch Automatisierung, vor allem für geringer qualifizierte Tätigkeiten, gerade auch im Bürobereich. Andererseits kann die Entlastung von monotonen Teilaufgaben auch Raum für kreativere Aufgaben schaffen. Die Überlegenheit der KI Systeme im Bereich der ärztlichen Diagnostik, im Bereich Umweltmanagement, um nur einige Anwendungen zu nennen, ist bereits vielfach nachgewiesen. Neue Tätigkeiten und Berufe werden entstehen, im Datenmanagement, im KI-Trainingsbereich. Insgesamt könnte es jedoch eine neue Polarisierung zwischen den qualifizierten Kräften, die die KI für sich nutzen und einsetzen geben und denen, die durch Nichtnutzung ihren „Marktwert“ verlieren.

Reflexionsfragen

- Welche konkreten Auswirkungen auf Arbeitsplätze und Tätigkeiten erwartest du in deinem Unternehmen durch KI und Automatisierung?
- Wie werden die Auswirkungen diskutiert?
- Welche Chancen bieten KI-unterstützte Assistenzsysteme für die Vereinfachung bestimmter Aufgaben? Welche neuen und kreativen Tätigkeiten könnten entstehen?
- Wie blickst du selbst auf die Entwicklung von KI? Siehst du eher Chancen oder Risiken für Deine Entwicklung. Wo möchtest du selbst von KI profitieren und sie einsetzen? Wo bist du zurückhaltend oder skeptisch?

Transparenz und Nachvollziehbarkeit

Diskutiert wird die Transparenz der (insbesondere kommerziellen) Modelle und Ihrer Trainingsalgorithmen und die Frage, woher die KI-Anwendungen die Basis für die generierten Inhalte beziehen. Die

⁷⁴<https://pictory.ai>

Transparenz von KI-Systemen ist entscheidend für Vertrauen und Verantwortlichkeit. Was passiert in der Black-Box zwischen einem Prompt und dem Output, insbesondere wenn auf der Basis der KI-Antworten automatisierte Entscheidungen getroffen werden?

Reflexionsfragen

- Wie stellt meine Organisation die Transparenz und Nachvollziehbarkeit von KI-Systemen sicher?
- Können wir die Entscheidungsprozesse unserer KI-Systeme verständlich erklären oder sind sie eine Black Box?
- Wie transparent machen wir die eingesetzten Algorithmus-Modelle und Trainingsdaten gegenüber Kunden und Nutzern?
- Welche Überwachungs- und Prüfsysteme gibt es, um Fehlentscheidungen von KI zu erkennen und zu korrigieren?
- Wie kommunizieren wir offen mit Kunden, wenn es doch zu Fehlern kommt?
- Schulen und trainieren wir unsere Mitarbeitenden, um KI-Systeme kompetent zu überwachen?
- Wie können wir als Unternehmen zu mehr Transparenz und Verständlichkeit von KI beitragen?

Verzerrungen und Diskriminierungen

KI-Systeme können bestehende Verzerrungen und Diskriminierungen widerspiegeln und verstärken, wenn die zugrundeliegenden Daten unfair sind oder Stereotypen enthalten. Häufig diskutiert wird der Einsatz von KI-Systemen bei der Bewerberauswahl oder im Finanz- und Versicherungsbereich, etwa bei der Kreditgewährung. Algorithmische Vorurteile in KI-Systemen können verschiedene Formen annehmen, wie Geschlechterbias, rassistische Vorurteile und Altersdiskriminierung.

Reflexionsfragen

- Enthalten die Daten, mit denen KI in unserem Unternehmen trainiert wird, möglicherweise versteckte Vorurteile und Verzerrungen?
- Spiegeln die Daten die Vielfalt der Gesellschaft wider oder nur kleine privilegierte Gruppen?
- Wie divers und interdisziplinär sind die Teams, die KI entwickeln?
- Welche Testverfahren gibt es, um Diskriminierungen in KI-Systemen aufzudecken und zu beseitigen?
- Wie kann mehr Bewusstsein für diese Problematik geschaffen werden?

Privatsphäre und Datenschutz

Der Einsatz von KI wirft eine Vielzahl von Fragen zum Umgang mit persönlichen Daten auf. Datenschutzverletzungen durch unsachgemäße Handhabung von KI-Systemen können gravierende Folgen haben.

Reflexionsfragen

- Welche persönlichen Kundendaten nutzen wir für unsere KI-Systeme? Sind die Daten richtig pseudonymisiert?
- Wie transparent machen wir die Verwendung von Kundendaten durch KI? Welche Zustimmungen holen wir ein?
- Wie stellen wir sicher, dass KI-Systeme Daten nicht unkontrolliert für unbeabsichtigte Zwecke verwenden? Welche Konsequenzen hätte es, wenn firmeninterne Daten in öffentlich zugängliche Systeme gelangen würden?
- Werden Datenschutz-Folgenabschätzungen vor dem Einsatz von KI-Systemen gemacht?
- Wie schulen und sensibilisieren wir unsere Mitarbeitenden zum sicheren und verantwortungsvollen Umgang mit Daten?

Digitale Abhängigkeit

Die KI hat das Potenzial, unsere kognitiven Fähigkeiten zu erweitern, Entscheidungsfindungen zu verbessern, birgt aber auch die Gefahr, Überabhängigkeit zu schaffen. Mit dem Einzug von KI in immer mehr Lebensbereiche wachsen die Sorgen mancher vor einer abnehmenden menschlichen Kompetenz und wachsender Abhängigkeit von der Technik. Vereinfacht gesagt: Werden wir schlauer oder dümmer durch KI? Werden wir durch das Verlassen auf die KI in gewissen Umfang entmündigt?

Reflexionsfragen

- Welche Fähigkeiten werden in einer von KI geprägten Arbeitswelt an Bedeutung gewinnen? Kreativität, Sozialkompetenz, Problemlösung,...
- Bieten wir Austauschforen, um Ängste vor KI abzubauen und Souveränität im Umgang mit ihr zu gewinnen?
- Bleibt der Mensch letzte Entscheidungsinstanz bei kritischen KI-Anwendungen oder überlassen wir wichtige Prozesse vollständig dem Algorithmus?
- Wie stärken wir die Medienkompetenz, um Fehlentwicklungen zu erkennen und ihnen entgegenzuwirken?

Ethik und Wertesystem

Die ethische Dimension der KI umfasst verschiedene Bedenken, wie Fairness und Verantwortlichkeit. Diskutiert wird, wem die KI dienen sollte: Dem Wohle aller Menschen und nicht nur weniger Konzerne. Welche Gefahren birgt die Monopolisierung? Und wer trägt die Verantwortung und ist rechenschaftspflichtig, insbesondere dann, wenn KI-Systeme - vielleicht sogar - eigenständig - fehlerhafte oder schädliche Entscheidungen treffen? Sollten Hersteller haften? Oder die Nutzer?

Reflexionsfragen

- Welche ethischen Richtlinien für KI gibt es in meiner Firma? Wer war an ihrer Erstellung beteiligt?
- Spiegeln die Richtlinien auch meine persönlichen Werte wider wie Gerechtigkeit, Verantwortung und Nachhaltigkeit?

- Sind Prozesse vorhanden, um ethische Fragen interdisziplinär zu diskutieren?
- Wie lässt sich die Einhaltung ethischer Prinzipien über den gesamten Entwicklungsprozess von KI-Systemen sicherstellen?
- Welche Schulungen braucht es, um Bewusstsein und Kompetenzen zu Ethik und KI zu stärken?

Regulierung

Das Spannungsfeld hier ist der Interessenausgleich zwischen dem Ausschöpfen der Innovationspotenziale und der Risikomimierung. Manche befürchten durch Regulierung eine Behinderung von Innovationen. Andere sehen Risiken für Gesellschaft und Demokratie, wenn KI völlig unreguliert eingesetzt wird. Wichtig ist in dem Zusammenhang auch die Diskussion um die politische Ebene, auf der eine mögliche Regulierung stattfinden kann, wenn sie wirksam sein sollte: National, europäisch, international oder branchenbezogen für besonders sensible Bereiche? Welche Rolle können in dem Zusammenhang die vielfach diskutierten freiwilligen Selbstverpflichtungen von Unternehmen oder Organisationen als Alternative zu Regelungen mit Sanktionen spielen?

Reflexionsfragen

- Wo könnten in meinem Unternehmen regulierungsbedürftige Risiken bestehen?
- Gibt es bereits interne Regeln oder Prinzipien für verantwortungsvolle KI in meiner Firma? Sollte das ausgebaut werden?
- Wie können hohe KI-Standards und Innovationsfähigkeit gleichzeitig gesichert werden?
- Sollte es breite gesellschaftliche Debatten zur Regulierung geben? Wie kann man sich konstruktiv einbringen?

2.9 Weiterführende Informationen und Links

Kurse:

- Künstliche Intelligenz und Maschinelles Lernen in der Praxis⁷⁵, Anbieter: OpenHPI
- Was bedeutet generative KI für unsere Gesellschaft?⁷⁶, Anbieter: OpenHPI
- Oxford Artificial Intelligence Programme⁷⁷

Webseiten & Blogs:

- ...

Podcasts, Videos:

⁷⁵<https://open.hpi.de/courses/kipraxis2021>

⁷⁶<https://open.hpi.de/courses/kizukunft2023>

⁷⁷<https://www.sbs.ox.ac.uk/programmes/executive-education/online-programmes/oxford-artificial-intelligence-programme>

- ...

Communitys:

- LAION⁷⁸

3 Lernpfad

Ein lernOS Lernpfad ist eine Zusammenstellung von Übungen (Katas), mit denen du neue Fähigkeiten erlernst und im Lauf der Zeit eine neue Haltung entwickelst. Ein Lernpfad kann innerhalb eines lernOS Sprints durchlaufen werden. lernOS Einsteiger*innen (NOOBs) können zum Start aus einem drei Lernpfaden “Zielorientierung & Vernetzung”, “Produktivität & Stressfreiheit” sowie “Offenheit & Vernetzung” wählen.

Die **Woche 0** ist die Woche, bevor es mit den Übungen im Lernpfad so richtig los geht. Nach Abschluss der Woche 0 sollten folgende Dinge festgelegt sein:

1. Welchen Lernpfad du verwendest
2. Ob du alleine, im Tandem oder im Circle lernst
3. Welche Hilfsmittel du zur Selbstorganisation verwendest (z.B. Videokonferenz-Tool für Circle)
4. Bei Tandem und Circle: An welchen Terminen die wöchentlichen Treffen stattfinden und wann du dir zusätzlich Lernzeit im Kalender reserviert hast.
5. Bei Tandem und Circle: ob ihr den Flipped-Classroom-Ansatz wählt und die Katas vor dem Weekly oder im Weekly macht.

Für die Woche 0 gibt es mit dem **Past Forward** und der Arbeit mit dem **lernOS Canvas** (siehe Kapitel in den Grundlagen) zwei zusätzliche Hilfsmittel. Der lernOS Canvas hilft, operative Felder Aufgaben, Projekte, Zielgruppen etc. zu visualisieren und daraus Ziele für diesen oder zukünftige Sprints abzuleiten. Mit dem Past Forward verschaffst du dir in kurzer Zeit Überblick über deine aktuelle Situation, wichtige Ereignisse der Vergangenheit und vor allem deine Vision für die Zukunft.

Anleitung Past Forward:

Das Past Forward basiert auf der Methode The Future, Backwards⁷⁹ von Dave Snowden. Mit dem Past Forward erhält man eine gute Sicht auf die persönliche Gesamtsituation durch einen Blick in die Vergangenheit und auf mögliche Zukünfte. Die Perspektive kann kurzfristig (1-2 Jahre), mittelfristig (3-5 Jahre) oder langfristig (ganzes Leben) sein. So gehst du vor:

⁷⁸<https://laion.ai>

⁷⁹<https://cognitive-edge.com/methods/the-future-backwards/>

1. Bereite deine Past Forward Dokumentation vor (Beispiel⁸⁰). Das kann im einfachsten Fall ein Blatt Papier im Querformat mit einem um 90 Grad nach rechts gedrehten “Y” darauf sein, Das Y stellt die aktuelle Situation (current state), die Vergangenheit, die Vision (heaven), die Anti-Vision (hell) sowie den “Stairway to Heaven” dar. Definiere die Zeitspanne, die du in Vergangenheit und Zukunft schauen möchtest (kurz-/mittel-/langfristig) (5 Minuten)
2. Beschreibe deine aktuelle Situation in 3-5 kurzen Sätzen (5 Minuten)
3. Beschreibe die 3-5 Schlüssel-Ereignisse in der Vergangenheit, die zur aktuellen Situation geführt haben (5 Minuten)
4. Beschreibe deine Vision in 3-5 kurzen Sätzen (5 Minuten)
5. Beschreibe deine Anti-Vision in 3-5 kurzen Sätzen (5 Minuten)
6. Beschreibe die 3-5 Schlüssel-Aktivitäten oder -Projekte, die deine Vision Wirklichkeit werden lässt und die Anti-Vision verhindert (5 Minuten)

4 Lernpfad

Idee: Ziel für den Lernpfad ist, mindestens einen Beitrag für Blog/Linkedin zu erstellen und zu veröffentlichen (Beispiel: mit Zusammenfassung Interview⁸¹). Das Ziel kann beliebig erweitert werden (mehrere Beiträge, mehrsprachig, Podcast, Video etc.)

4.1 Woche 0

Die erste Woche dient dem Kennenlernen. Hier findest Du Woche 0⁸².

4.2 Woche 1

Lese Dir den Abschnitt aus dem Grundlagenkapitel KI & Gesellschaft⁸³ durch und überlege Dir wie eine verantwortungsvolle Nutzung von KI-Systemen aussehen könnte.

Beantworte diese Fragen vorab:

- Überlege Dir, welche Eingaben und Ausgaben es bei generativen KIs gibt und welche Daten generative KIs nutzen können, z. B. für das Training.
- Was wären Bedrohungsszenarien, welche durch eine verantwortungslose Nutzung von generativer KI für Dein Unternehmen entstehen könnten?

⁸⁰https://cognitive-edge.com/wp-content/uploads/2015/01/3---ChrisFI-IMG-0058-wpcf_300x225.jpg

⁸¹<https://www.linkedin.com/feed/update/urn:li:activity:7110171493103198209/>

⁸²2-0-Woche-0.md

⁸³1-7-ai-and-society.md

- Überlege Dir, was bei der Nutzung von KI-Systemen zu beachten ist. Informiere Dich im Intranet Deines Unternehmens, ob es bei Dir unternehmensspezifische Vorgaben und Richtlinien zur Nutzung von KI-Systemen im Internet gibt.

Für das Treffen in der Gruppe:

- Erstellt als Gruppe einen kurzen Steckbrief darüber, was man aus Eurer Sicht bei einer verantwortungsbewusster Nutzung von KI beachten sollte.

4.3 Woche 2

- Wähle einen KI-Chatbot aus. Als Hilfestellung kannst Du Kapitel 1.6⁸⁴ nutzen.
 - nutze das Dir bekannte KI-Tool, dass im Unternehmen erlaubt ist
 - Unsere Empfehlungen wäre ChatGPT oder sowas wie LLAMA 2 auf Perplexity Labs
- Besorge Dir einen Zugang
- Erste Eingabe: “Ich packe meinen ...” -> Koffer
- Einstiegs-Scherzfragen: KI-Einstieg mit ChatGPT-Scherzfragen | eBildungslabor⁸⁵

4.4 Woche 3

Lese Dir das Kapitel 1.2⁸⁶ durch + Übung dazu.

- Was ist der Unterschied zwischen “Maschinellern Lernen” und “KI”?
 - Hinweis: Grundlagenkapitel unvollständig. Siehe Unterscheidung von AI > ML > DL.
 - Eine gute Vorlage könnte das hier sein: Glossar: Acht KI-Begriffe, die Ihnen auch im Alltag begegnen (handelsblatt.com)⁸⁷
- Lass Dir von einer KI erklären, wie ein KI-Chatbot funktioniert
- Überlege Dir, was für ein KI-Lernprojekt Du machen willst
 - Entscheide Dich für ein Medium (Text, Bild, ...)
 - es muss etwas sein, dass von der KI generiert wird

⁸⁴1-6-ai-tools-services.md

⁸⁵<https://ebildungslabor.de/blog/ki-einstieg-mit-chatgpt-scherzfragen/>

⁸⁶1-2-ai-machine-learning.md

⁸⁷<https://www.handelsblatt.com/technik/ki/glossar-acht-ki-begriffe-die-ihnen-auch-im-alltag-begegnen/29431606.html>

4.5 Woche 4

Boxenstopp 1

4.6 Woche 5

4.7 Woche 6

4.8 Woche 7

4.9 Woche 8

Boxenstopp 2

4.10 Woche 9

4.11 Woche 10

4.12 Woche 11

5 Anhang

5.1 Danksagungen

Ein großer Dank an die vielen Projekte und Vorarbeiten, die die Erstellung und Verwaltung von lernOS Leitfäden ermöglichen.

5.2 Änderungshistorie

Bearbeitet			
Version	von	Beschreibung Änderung	Datum
0.1	Simon Dückert	Erste Version des Leitfadens angelegt	22.09.2023

5.3 lernOS Glossar

A

Aufgabe Definition fehlt

Arbeitsumgebung Nach DIN 33400 ist die Arbeitsumgebung eines Arbeitssystems „das räumliche Umfeld, von dem vor allem physikalische und chemische, aber auch unter anderem biologische (z. B. bakteriologische) Einflüsse auf den Menschen einwirken“. Hierbei wirken Mensch und Arbeitsmittel im Arbeitsablauf am Arbeitsplatz in einer Arbeitsumgebung unter den Bedingungen dieses Arbeitssystems zusammen.

C

Change Agent Definition fehlt

Change Management Definition fehlt

E

E-Portfolio Ein E-Portfolio ist eine persönliche, strukturierte und digitale Informationssammlung, die Lernziele und -prozesse im lebenslangen Lernen unterstützt und dabei erworbenes Wissen und Fähigkeiten veranschaulicht.

F

Formale Organisation Definition fehlt

Formales Lernen Formales Lernen ist das Lernen, das üblicherweise in einer Bildungs- oder Ausbildungseinrichtung stattfindet, in Bezug auf Lernziele, Lernzeit oder Lernförderung strukturiert ist und zur Zertifizierung führt. Formales Lernen ist aus der Sicht des Lernenden zielgerichtet (Europäische Kommission, 2001).

Führung Führung ist die Tätigkeit, Menschen in die Zielfindung einzubinden, motivierende Rahmenbedingungen zu schaffen, Lern- und Entwicklungsprozesse zu unterstützen und die Zielerreichung bei Bedarf zu unterstützen.

Führungskraft Eine Führungskraft ist eine Person in einer Organisation, die mit Aufgaben der Personalführung betraut ist.

Führungsstil Der Führungsstil bezeichnet das langfristig stabile und von der Situation unabhängige Verhalten einer Führungskraft gegenüber Mitarbeitern. Zu den klassischen Führungsstilen nach Levin gehören autoritäre Führung, demokratische Führung und Laissez-faire Führung.

G

Getting Things Done (GTD) Definition fehlt

I

Informale Organisation Definition fehlt

Informelles Lernen Informelles Lernen ist das Lernen, das im Alltag, am Arbeitsplatz, im Familienkreis oder in der Freizeit stattfindet. Es ist in Bezug auf Lernziele, Lernzeit oder Lernförderung nicht strukturiert und führt üblicherweise nicht zur Zertifizierung. Informelles Lernen kann zielgerichtet sein, ist jedoch in den meisten Fällen nichtintentional beiläufig (Europäische Kommission, 2001)

Infrastruktur Definition fehlt

K

Kollaboration Kollaboration ist die Zusammenarbeit von Personen oder Gruppen von Personen (Wikipedia).

Kommunikation Kommunikation ist der Austausch oder die Übertragung von Informationen, die auf verschiedenen Arten (verbal, nonverbal) oder verschiedenen Wegen (Sprechen, Schreiben) stattfinden kann (Wikipedia).

Kompetenz ability to apply knowledge and skills to achieve intended results (ISO 9001:2015).

L

Lebenslanges Lernen Lebenslanges Lernen bezeichnet alles Lernen während des gesamten Lebens, das der Verbesserung von Wissen, Qualifikationen und Kompetenzen dient und im Rahmen einer persönlichen, bürgergesellschaftlichen, sozialen, bzw. beschäftigungsbezogenen Perspektive erfolgt (Europäische Kommission, 2001).

Lernen Lernen ist der absichtliche oder beiläufige Erwerb von Wissen und Fähigkeiten. Lernen führt zu einer Veränderung des Verhaltens, Denkens oder Fühlens auf Basis neuer Erfahrungen oder Einsichten.

Lernende Organisation Eine Lernende Organisation ist eine Organisation, die die Fähigkeit besitzt, Wissen zu generieren, zu akquirieren und zu verteilen und ihr Verhalten auf Basis neuer Erkenntnisse und Einsichten zu verändern.

lernOS lernOS ist ein offenes System für Lebenslanges Lernen und Lernende Organisationen. Es stellt auf den drei Ebenen Individuum, Team und Organisation offene Leitfäden für die kontinuierliche Verbesserung des Lern- und Wissensmanagements bereit.

lernOS Canvas Definition fehlt

lernOS Rad Definition fehlt

Lernpfad Definition fehlt

Lernsprint Definition fehlt

Lernzirkel Eine Gruppe von 4-5 Personen in der sich die Mitglieder gegenseitig mit Feedback, Erfahrung, Wissen und Reflexion helfen. Die Circle-Mitglieder treffen sich wöchentlich und folgen dabei einem vorgegebenen Ablauf, der den Lern- und Entwicklungsprozess strukturiert.

M

Management Management ist die Koordination der Aktivitäten in einer Organisation mit dem Zweck, die Ziele der Organisation zu erreichen.

Mindset Definition fehlt

Mission Die Mission beschreibt, was die Organisation im Tagesgeschäft tut oder tun wird (Business Motivation Model).

O

Objectives & Key Results (OKR) Definition fehlt

Organisation Eine Organisation besteht aus einer oder mehreren Personen (Organisationsmitglieder), die ein gemeinsamer Zweck verbindet und die sich zur Zielerreichung eine formale Organisationsstruktur geben.

P

Projekt Definition fehlt

Prozess set of interrelated or interacting activities that use inputs to deliver an intended result (ISO 9001:2015).

R

Rolle Definition fehlt

S

Selbstgesteuertes Lernen Definition fehlt

Selbstorganisation Definition fehlt

Sinn Definition fehlt

Skillset Definition fehlt

Sprint Definition fehlt

Strategie Definition fehlt

T

Toolset Definition fehlt

True North Der "wahre Norden" (true north) ist das langfristige Ziel (auch Vision genannt), an dem sich jährliche Ziele und konkrete Verbesserungsaktivitäten orientieren.

V

Vision Die Vision beschreibt den zukünftigen Zustand einer Organisation und ihrer Umwelt, unabhängig davon, wie dieser erreicht wird (Business Motivation Model).

VUCA Die Abkürzung steht für volatility, uncertainty, complexity und ambiguity.

W

Wissen Definition fehlt

Wissensmanagement Wissensmanagement ist die Führung und Gestaltung einer Lernenden Organisation.

Working Out Loud Working Out Loud (WOL) ist eine Arbeitsweise, bei der Arbeitsstände und -ergebnisse offen einsehbar sind (z.B. Wiki, offene Dateiablage) und über den Arbeitsfortschritt offen erzählt wird (z.B. Blog, Microblog).

Working Out Loud Lernprogramm Das Working Out Loud Lernprogramm von John Stepper ist ein 12-wöchiger Lernpfad, mit dem Einsteiger*innen die WOL Arbeitsweise erlernen können.

Z

Ziel Result to be achieved (ISO 9001:2015).