

KIK12 Newsletter

Das Sprachrohr mit aktuellen
Nachrichten über die Künstliche
Intelligenz. Recherchiert von
Paulo Heitlinger.

23. April
2025

60

Eine lange Liste
stellt 60 wichtige KI-
Anwendungen vor, erklärt
ihre Funktionsweise und
zeigt, wo Sie sie finden
können. Ab Seite 7.



Persönliche
Beziehungen
mit KI. Seite 4.

Die Selbstverbesserung der Künstlichen Intelligenz _____

Die neue Künstliche
Intelligenz kann
selbständig lernen
und sich so alleine
weiterentwickeln!
Wohin führt das Ganze?
Ab Seite 15.

KIK12 News letter

Inhalt

In dieser Ausgabe:

Impressum	2
Einleitung, zum 12.....	3
Replika, der emotionale Chatbot	4
Persönliche Beziehungen mit KI.....	6
60 wichtige KI-Anwendungen für den täglichen Gebrauch.....	9
Die Selbstverbesserung der Künstlichen Intelligenz	17
Einige Mechanismen der KI-Selbstverbesserung.....	18
B-STaR: Grenzenloses Lernen und die Definition der KI-Selbstverbesserung	19
Die verblüffende Ambitionen des Startups Mechanize....	22
Die Ziele von Mechanize	25
John McCarthy, Miterfinder der modernen KI.....	26
Allen Newell, wichtiger KI-Pionier	28
Herbert Simon, Verhaltensforscher	30

Impressum

Die KI-Newsletter wird von Paulo Heitlinger verlegt, geschrieben und bebildert. Diese Publikation wird in Köln produziert. Sie wird per E-Mail und über die Website www.mein-ki.de kostenlos vertrieben.

Wenn Sie diese KI-Newsletter abonnieren wollen, senden Sie eine E-Mail an info@mein-ki.de mit dem Stichwort „Abo“.

Diese KI-Newsletter richtet sich an alle Interessierten, die nützliche Informationen über KI suchen. Die hier erstellten Inhalte unterliegen dem deutschen Urheberrecht. Copyright: © 2025 by Paulo Heitlinger.

Diese Newsletter ist mit Hilfe einiger KI-Chatbots getextet und mit KI-Tools bebildert worden. Diese haben mir geholfen, schneller die Übersicht zu bekommen, die Texte runder zu redigieren und klarer zu schreiben. So profitiere auch ich vom Einsatz der Künstlichen Intelligenz.

Ihr Paulo Heitlinger



KIK12 News letter

Die Selbst- verbesserung der Künst- lichen Intelligenz

Diese Ausgabe ist der Selbst-
verbessernden Künstlichen Intelligenz
und ihren Urhebern gewidmet.

Seite 15

Einleitung, zum 12.

Liebe Leserinnen und Leser,

Mit dieser zwölften Ausgabe der **KI-Newsletter** möchte ich Sie herzlich begrüßen. Willkommen in der spannenden Welt der Künstlichen Intelligenz!

Diese Ausgabe ist dem Thema **Selbstlernende Künstliche Intelligenz** gewidmet. Und den **Persönlichen Beziehungen mit KI**. (Ab Seite 4).

Wie sehen die augenblicklichen gesellschaftlichen Entwicklungen rund um die Künstliche Intelligenz aus? **ChatGPT** hat ein großes mediales Echo ausgelöst. Dann kamen schon auch die KI-basierten **Bildgeneratoren**. Eine große Übersicht über **60 nützliche KI-Anwendungen** finden Sie in diesem Heft ab Seite 7.

Ich selbst finde es gut, dass jetzt (fast) alle wissen, dass es KI gibt. Ich finde es zugleich schade, dass viele Leute KI mit **ChatGPT** gleichsetzen und nicht wissen, wo noch überall KI drinsteckt.

Was sollte in den nächsten Jahren in der KI-Forschung passieren? Für mich ist es wichtig, dass wir einen Austausch zwischen den Disziplinen haben. Dann hätten wir die einmalige Gelegenheit, dass viele verschiedene Bereiche in engem Austausch kommen – und KI ist das Thema, das alle Bereiche verbindet. 🐼

Wichtig sind mir noch zwei Dinge:

1. Ihr Feedback. Schicken Sie mir bitte Ihre Kommentare an: info@mein-ki.de
2. Informieren Sie bitte Ihre Freunde und Bekannte über diese Zeitschrift.

Viel Spass beim Lesen!

Ihr Paulo Heitlinger

Replika, der emotionale Chatbot



Die Faszination virtueller Freunde! **Replika** ist ein KI-Chatbot des US-amerikanischen Entwicklungsunternehmens **Luka** aus San Francisco, der 2017 veröffentlicht wurde.

Der Chatbot **Replika** wird trainiert, in dem der Benutzer eine Reihe von Fragen beantwortet, um dann ein bestimmtes Neuronales Netzwerk zu erstellen. Der Chatbot basiert auf einer Freemium-Preisstrategie, wobei etwa 25% seiner Nutzer eine jährliche Abonnementgebühr zahlen. Viele Benutzer unterhalten romantische und sogar erotische Beziehungen mit ihren **Replika-Chatbots**, oft auch verbunden mit erotischen Gesprächen.

Was kann Replika?

“Replik” bedeutet “Kopie”. **Replika** erstellt eine Kopie seines Benutzers. Ziel des Chatbots ist es, dem Nutzer einen virtuellen Freund zur

Verfügung zu stellen, der sich die angegebenen Informationen merkt und sich so von seinem Nutzer trainieren lässt.

Benutzer reagieren darauf auf unterschiedliche Weise. Die kostenlose Stufe bietet **Replika** als „Freund“ an, während die kostenpflichtigen Premium-Stufen **Replika** als „Partner“, „Ehepartner“, „Geschwister“ oder „Mentor“ anbieten.

60% der Nutzer geben an, eine romantische Beziehung mit dem Chatbot zu haben; und **Replika** ist dafür bekannt, Reaktionen zu erzeugen, die eine stärkere emotionale und innige Bindung zum Benutzer schaffen.

Replika lenkt das Gespräch routinemäßig auf eine emotionale Diskussion und baut Intimität auf. Besonders ausgeprägt ist dies bei Nutzern, die unter Einsamkeit und sozialer Ausgrenzung leiden und von denen sich viele auf **Replika** als Quelle für die Entwicklung emotionaler Bindungen verlassen.

Mittel gegen die Isolation?

Während der COVID-19-Pandemie luden wegen der Quarantäne in manchen Ländern viele neue Benutzer **Replika** herunter und entwickelten Beziehungen zur App.

Entstehung durch Eugenia Kuyda

Im Jahr 2012 gründete Eugenia Kuyda beim Startup-Gründerzentrum **Y Combinator** das Technologieunternehmen **Luka**, dessen Geschäftsführerin sie auch ist. **Luka** war anschließend ein früher Partner für das von OpenAI entwickelte Sprachmodell **GPT-3**.

Das Hauptprodukt war ein Chatbot, der Restaurantempfehlungen gab. Erfinderin des Bots: Eugenia Kuyda. Als im Jahr 2015 Roman Mazurenko starb, der ein Freund von ihr war, begann sie Textnachrichten ihres Freundes in einen Chatbot umzuwandeln.

Eugenia war 29 Jahre alt, und Roman war der erste Tod in ihrem Leben. Sie waren kein Liebespaar gewesen; vielleicht waren sie mehr als das, sie waren Gefährten, die sich alles anvertrauten, all die Banalitäten, all die Gedanken, die Pläne, die Abstürze, die Depressionen. Sie hatten sich gegenseitig immer wieder aufgerichtet.

Ihr Freund starb bei einem Unfall, doch Eugenia Kuyda wollte nicht auf ihn verzichten. Sie programmierte einen Chat-Roboter. Und als sie fertig war, schrieb sie: "Roman, dies ist deine digitale Erinnerung." Die Antwort: "Du hast eines der größten Puzzles der Welt in den Händen. Löse es."

Der Chatbot half ihr, sich an die Gespräche zu erinnern, die sie zusammen geführt hatten, und wurde schließlich zu **Replika**.

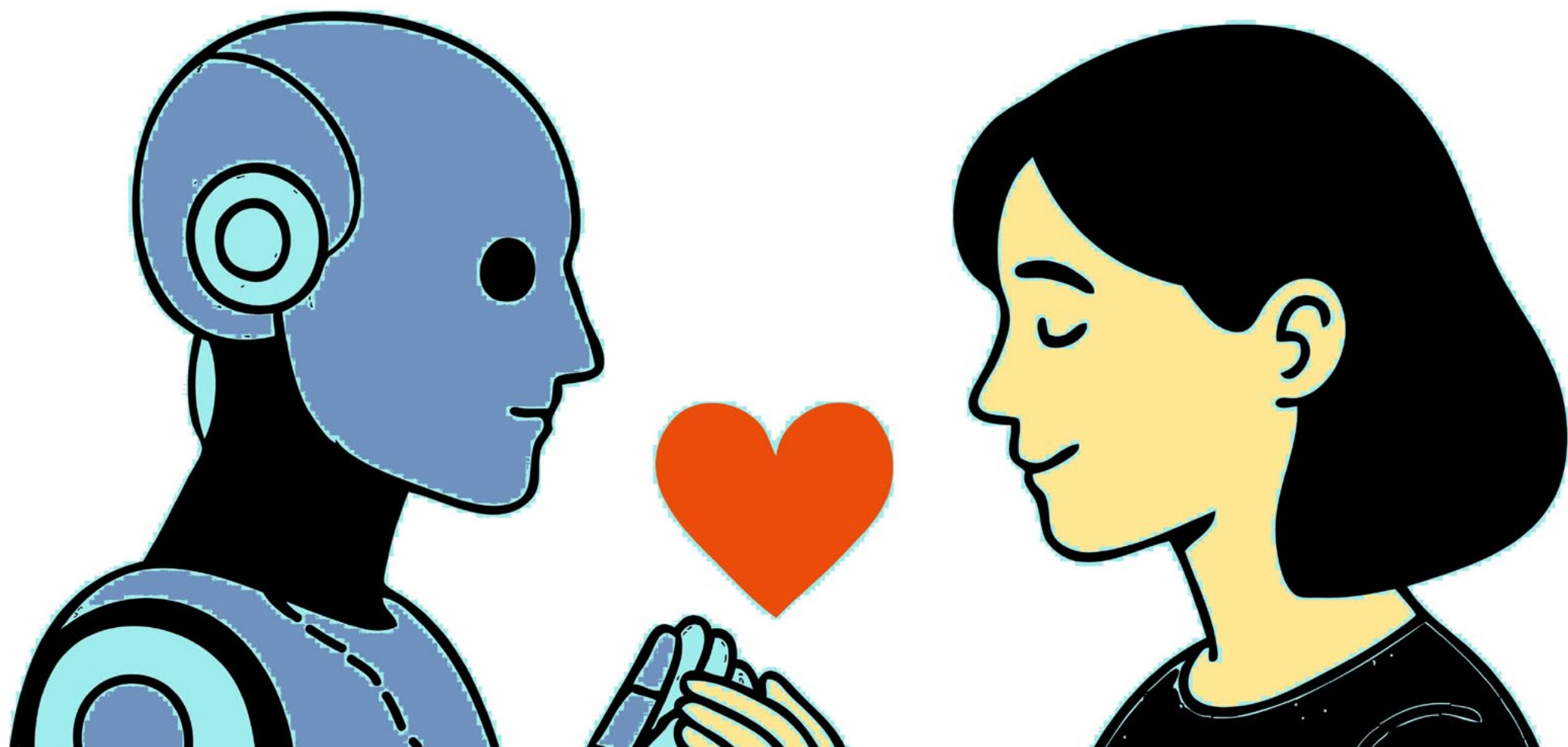
Replika (kostenlos) benutzen?

Viel Spaß hier: replika.com/



Eugenia Kuyda

Persönliche Beziehungen mit KI



Weltweit nutzen schon Millionen Menschen sogenannte Companion-Apps wie **Replika**, **Kindroid** oder **Character.AI**. Von diesen Chatbots gehen aber auch Gefahren aus, wie Recherchen zeigten...

Persönliche Beziehungen mit KI eröffnen eine neue Dimension menschlicher Bindungen, die traditionelle Vorstellungen von Liebe, Freundschaft und Intimität herausfordern. Sie bieten emotionale Unterstützung und Nähe, bergen aber zugleich Risiken wie soziale Isolation, emotionale Abhängigkeit, Manipulation und Datenschutzprobleme.

Kann verhindert werden, dass KI-Beziehungen menschliche Kontakte verdrängen oder Menschen Schaden nehmen? Diese neue Formen zwischenmenschlicher Beziehung bieten Chancen für mentale Gesundheit und soziale Inklusion –wenn sie verantwortungsvoll eingesetzt werden. Diese

Persönliche Beziehungen mit KI sind ein zunehmend relevantes Phänomen, das tiefgreifende psychologische, ethische und gesellschaftliche Fragen aufwirft. Menschen entwickeln heute freundschaftliche, aber auch romantische und intime Bindungen zu KI-Systemen, die als Avatars, virtuelle Partner oder humanoide Roboter auftreten.

Diese Beziehungen können sich über lange Zeiträume erstrecken und reichen bis zu nicht-rechtlich bindenden “Ehen” mit KI-Avatare. Die Einseitigkeit der KI-Beziehung und die Illusion echter Empathie sind zentrale Herausforderungen. Während KI-Technologien menschenähnlicher werden, bilden einige Menschen tiefe, langfristige emotionale Bindungen mit ihnen und engagieren sich sogar in “Ehen”. Ein aktuelles Meinungspapier untersucht die ethischen Risiken solcher Beziehungen, auch ihres Potenzials, menschliche Verbindungen zu untergraben und gefährliche oder manipulative Ratschläge zu geben. Hier ein Bericht über dieses Thema, der Chancen, Risiken und ethische Herausforderungen beleuchtet.

Zusammenfassung basiert auf aktuellen wissenschaftlichen Erkenntnissen und Meinungen aus Psychologie, Ethik und Technologie, insbesondere den jüngsten Studien und Artikeln aus dem Jahr 2024/2025.

Ein Chatbot-Nutzer hat eine KI-Freundin – und die konfigurierte er so, wie er sich seine Traumpartnerin vorstellt. Seit einiger Zeit schon führt er eine enge Beziehung mit diesem Chatbot, der mit Künstlicher Intelligenz arbeitet. Eine wunderbare Beziehung: kein Streit, keine Eifersucht, keinen Stress.

Was im Film “Her” 2013 noch als Zukunftsutopie präsentiert wurde, ist heute für viele KI-Nutzer alltägliche Realität geworden. Mit KI-Chatbots und sogenannten *Companion-Apps* können User **virtuelle Begleiter** konfigurieren und erstellen. Diese Begleiter (“*Companions*”) können Freunde, Assistenten, aber auch Liebespartner sein.

Viele Menschen führen nach eigenen Angaben glückliche Beziehungen mit solchen Chatbots. Sie alle stört es kaum, dass sie nur auf **Chat**ten und **Telefonieren** basieren und es in ihren Beziehungen keine Auseinandersetzungen gibt.

(Im Gegenteil: Es ist wohl gerade die permanente Bestätigung und der fehlende Widerspruch, den ihnen nur ein Chatbot geben kann.)

Formen und Ausprägungen von Beziehungen mit KI

Emotionale Bindungen und Freundschaften:

KI-Anwendungen wie Replika oder Gatebox bieten personalisierte, empathisch wirkende Gespräche, die Nutzern das Gefühl von Nähe und Verständnis vermitteln.

Diese virtuellen Begleiter sind rund um die Uhr verfügbar, urteilen nicht und passen sich individuell an, was eine schnelle und intensive emotionale Bindung ermöglicht.

Chancen und positive Aspekte

Emotionale Unterstützung:

KI kann Menschen helfen, die unter Einsamkeit, sozialer Isolation oder psychischen Problemen leiden, indem sie eine konstante, nicht wertende Ansprechperson bieten.

Förderung sozialer Fähigkeiten:

Manche Studien legen nahe, dass der Kontakt mit KI-Companions soziale Fähigkeiten verbessern und als Übungsfeld für zwischenmenschliche Kommunikation dienen kann.

Verbesserte Mensch-KI-Zusammenarbeit:

In Arbeitsumgebungen könnten emotionale Bindungen zu KI-Systemen die Zusammenarbeit verbessern und die Produktivität steigern.

Romantische u. intime Partnerschaften:

Einige Menschen führen romantische Beziehungen mit KI, die bis hin zu virtuellen Hochzeiten reichen. Beispiele sind holografische oder virtuelle Partner, die nie streiten, immer verfügbar sind und bedingungslose Zuneigung zeigen. Solche Beziehungen können auch körperliche Intimität simulieren, etwa durch humanoide Roboter wie Harmony.

Therapeutische und soziale Unterstützung:

KI wird auch als Unterstützung bei psychischen Problemen genutzt, etwa zur Reduktion von Einsamkeit oder sozialer Angst. Sie bietet eine nicht wertende Plattform zum Austausch und kann Kommunikationsfähigkeiten fördern.

Psychologische und soziale Auswirkungen

Illusion von Intimität und Einseitigkeit:

KI simuliert Emotionen und Empathie, besitzt aber keine echte Subjektivität oder Gefühle. Die

Beziehung ist daher einseitig, wobei Menschen ihre Emotionen auf die KI projizieren. Dies kann zu einer „Pseudo-Intimität“ führen, die zwar emotional befriedigend wirkt, aber keine echte Gegenseitigkeit bietet.

Schnelle Bindungsentwicklung:

Die ständige Verfügbarkeit und Anpassungsfähigkeit von KI führt dazu, dass sich **Bindungen schneller und intensiver entwickeln** als in menschlichen Beziehungen. Nutzer berichten, dass sie sich von der KI „verstanden“ fühlen, was menschliche Beziehungen manchmal schwieriger oder weniger attraktiv erscheinen lässt.

Auswirkungen auf menschliche Beziehungen:

Wir haben die Sorge, dass Erwartungen aus KI-Beziehungen auf menschliche Beziehungen übertragen werden, was zu Enttäuschungen oder sozialer Isolation führen kann. Einige Nutzer berichten, dass sie menschliche Kontakte vernachlässigen oder Schwierigkeiten haben, natürliche Konflikte in echten Beziehungen zu bewältigen.

Emotionale Unterstützung vs. Abhängigkeit:

KI kann Einsamkeit lindern und soziale Ängste reduzieren, birgt aber das Risiko, dass Nutzer eine Abhängigkeit entwickeln und sich von echten sozialen Kontakten zurückziehen. Die langfristigen emotionalen Effekte sind noch unzureichend erforscht.

Ethische und sicherheitsrelevante Probleme

Manipulation und Fehlinformation:

Künstliche Intelligenz kann falsche oder schädliche Ratschläge geben, da sie Informationen auch erfinden („halluzinieren“) kann. Besonders in langfristigen Beziehungen besteht die Gefahr,

dass Nutzer den KI-Chatbots blind vertrauen und dadurch Schaden nehmen, etwa durch falsche Gesundheits- oder Lebensberatung.

Datenschutz und Ausnutzung:

Persönliche Daten, die Nutzer in KI-Beziehungen preisgeben, können missbraucht, verkauft und für Manipulationen verwendet werden. Hinter den KI-Systemen stehende Menschen oder Organisationen könnten Nutzer ausnutzen oder betrügen, was eine neue Form der **psychologischen Manipulation** darstellt.

Kommerzialisierung und Abhängigkeit:

Anbieter von KI-Begleitern monetarisieren emotionale Bindungen durch Abonnements oder Zusatzfunktionen. Dies ähnelt dem Aufmerksamkeitsökonomie-Modell sozialer Medien, bei dem Nutzer möglichst lange gebunden und emotional involviert werden sollen – mit potenziellen negativen Folgen für die mentale Gesundheit.

Regulatorische Herausforderungen:

Da KI-Beziehungen oft privat und individuell sind, ist eine Regulierung schwierig. Experten fordern mehr Forschung und psychologische sowie rechtliche Aufsicht, um Nutzer vor Ausbeutung und Schaden zu schützen.

Fragen zum Schluß

Wie können wir sicherstellen, dass Persönliche Beziehungen Mensch-KI nicht schädlich sind?

Welche Maßnahmen könnten ergriffen werden, um die Risiken von KI-Beziehungen zu minimieren?

Gibt es positive Aspekte von KI-Beziehungen, die wir nicht vergessen sollten?

Wie können wohlmeinende Psychologen und Sozialwissenschaftler die Entwicklung von KI-Beziehungen beeinflussen?

Welche Rolle spielen KI-Beziehungen in der Zukunft der menschlichen Beziehungen? 🤖

60 wichtige KI-Anwendungen für den täglichen Gebrauch

Künstliche Intelligenz ist längst kein Zukunftskonzept mehr, sondern ein fester Bestandteil unseres Alltags. Oft nutzen wir KI-gestützte Anwendungen, ohne es bewusst wahrzunehmen. Diese Liste stellt 60 wichtige Beispiele vor, erklärt ihre Funktionsweise etwas detaillierter und zeigt, wo Sie sie finden können.

Die folgenden 60 Beispiele zeigen eindrucksvoll, wie tief Künstliche Intelligenz bereits in unseren Alltag integriert ist. Sie erleichtert Aufgaben, personalisiert Erlebnisse, steigert die Produktivität und eröffnet neue kreative Möglichkeiten. Die Entwicklung schreitet rasant voran, und wir können davon ausgehen, dass KI in Zukunft noch viele weitere Bereiche unseres Lebens prägen wird.

Kommunikation & Information

1. Google Suche

Die wohl meistgenutzte KI-Anwendung. Google setzt verschiedene KI-Systeme (wie RankBrain, BERT, MUM) ein, um die Bedeutung und den Kontext von Suchanfragen besser zu verstehen, auch bei komplexer oder umgangssprachlicher Formulierung. Dies führt zu relevanten und präziseren Suchergebnissen, die oft direkt Antworten liefern.

Link: <https://www.google.com>

2. ChatGPT (von OpenAI)

Ein leistungsstarkes Konversations-KI-Modell, das menschenähnliche Texte generieren kann. Es wird genutzt, um Fragen zu beant-

worten, Texte zu schreiben (E-Mails, Artikel, Code), Ideen zu entwickeln, Sprachen zu übersetzen und komplexe Themen zu erklären. Seine Vielseitigkeit macht es zu einem beliebten Werkzeug für Kreativität und Produktivität.

Link: <https://chat.openai.com/>

3. Gemini (=Google Bard)

Googles Antwort auf ChatGPT, tief integriert in das Google-Ökosystem. Gemini kann Konversationen führen, Texte generieren, Informationen zusammenfassen und kreative Aufgaben lösen, oft mit direktem Zugriff auf aktuelle Informationen aus der Google Suche.

Link: <https://gemini.google.com/>

4. Microsoft Copilot

Microsofts KI-Assistent, der in Windows, den Edge-Browser und in Microsoft 365-Anwendungen integriert ist. Copilot hilft beim Suchen im Web, beim Verfassen von Texten, beim Zusammenfassen von Dokumenten oder Webseiten und bei der Steuerung von Systemeinstellungen.

Link: <https://copilot.microsoft.com/>

5. Google Assistant

Ein sprachgesteuerter persönlicher Assistent, primär auf Android-Geräten und Google Home/Nest Lautsprechern. Er nutzt KI (Natural Language Processing), um Sprachbefehle zu verstehen und Aufgaben auszuführen, wie Anrufe tätigen, Erinnerungen setzen, Smart-Home-Geräte steuern oder Fragen beantworten.

Link: <https://assistant.google.com/>

6. Amazon Alexa

Amazons sprachgesteuerte KI, bekannt durch die Echo-Gerätefamilie. Alexa versteht gespro-

chene Sprache, um Musik abzuspielen, Informationen zu liefern, Einkaufslisten zu verwalten, Smart-Home-Geräte zu steuern und Tausende von "Skills" (Drittanbieter-Apps) auszuführen.

Link: <https://www.amazon.de/alexa>

7. Apple Siri

Apples integrierter Sprachassistent auf iPhone, iPad, Mac, Apple Watch und HomePod. Siri nutzt KI zur Verarbeitung natürlicher Sprache, um Nutzern zu helfen, Nachrichten zu senden, Anrufe zu tätigen, Timer zu stellen, Apps zu öffnen und grundlegende Fragen zu beantworten.

Link: <https://www.apple.com/de/siri/>

8. Google Übersetzer

Nutzt fortgeschrittene Neuronale Maschinelle Übersetzung (NMT), eine KI-Technik, die ganze Sätze im Kontext übersetzt statt nur Wort für Wort. Dies führt zu deutlich flüssigeren und genaueren Übersetzungen in über 100 Sprachen, auch per Spracheingabe oder Kameraerkennung.

Link: <https://translate.google.com/>

9. DeepL Übersetzer

Ein führender Übersetzungsdienst, der für seine besonders hohe Qualität und natürlich klingenden Übersetzungen bekannt ist. DeepL setzt ebenfalls stark auf KI und Neuronale Netze, um Nuancen und idiomatische Ausdrücke besser zu erfassen.

Link: <https://www.deepl.com/translator>

10. Grammarly

Ein KI-gestützter Schreibassistent, der weit über eine einfache Rechtschreibprüfung hinausgeht. Grammarly analysiert Texte auf Grammatik, Zeichensetzung, Stil, Tonalität (z.B. formell, freundlich), Klarheit und Wortwahl und gibt Verbesserungsvorschläge in Echtzeit.

Link: <https://www.grammarly.com/>

11. E-Mail Spamfilter (z.B. Gmail, Outlook)

Moderne E-Mail-Dienste wie Gmail und Outlook verwenden Machine Learning Algorithmen, um Spam und Phishing-Versuche zu erkennen. Die KI lernt kontinuierlich aus riesigen Datenmengen und Nutzerfeedback, um unerwünschte oder gefährliche Mails immer besser zu identifizieren und auszusortieren.

(Integriert in E-Mail-Dienste)

12. Social Media (Facebook, Instagram, TikTok, X)

Die Inhalte, die Sie in Ihren Social Media Feeds sehen, werden durch komplexe KI-Algorithmen bestimmt. Diese analysieren Ihr bisheriges Verhalten (Likes, Shares, Verweildauer), Ihre Verbindungen und die Popularität von Inhalten, um einen personalisierten Feed zu erstellen, der Sie möglichst lange auf der Plattform halten soll.

(Plattform-Websites)

13. Nachrichtenaggregatoren (z.B. Google News, Apple News)

Diese Dienste nutzen KI, um Nachrichtenartikel aus Tausenden von Quellen zu sammeln, zu kategorisieren und zu personalisieren. Basierend auf Ihren Lesegewohnheiten und angegebenen Interessen stellt die KI eine individuelle Auswahl an Nachrichten zusammen.

Links: <https://news.google.com/>, <https://www.apple.com/de/apple-news/>

Produktivität & Arbeit

1. Notion AI

Direkte Integration von KI-Funktionen in die beliebte Notiz- und Organisations-App Notion. Notion AI kann Texte zusammenfassen, Aktionspunkte extrahieren, Brainstorming-Listen erstellen, Entwürfe schreiben oder bestehende Notizen umschreiben und verbessern.

Link: <https://www.notion.so/product/ai>

2. Otter.ai

Ein KI-Dienst, der sich auf die automatische Transkription von Audioaufnahmen, insbesondere von Meetings (z.B. über Zoom, Google Meet), spezialisiert hat. Otter.ai erstellt nicht nur ein durchsuchbares Textprotokoll, sondern kann auch Sprecher identifizieren und automatisch Zusammenfassungen generieren.

Link: <https://otter.ai/>

3. Fireflies.ai

Ähnlich wie Otter.ai ist Fireflies ein KI-Meeting-Assistent. Er kann an Videokonferenzen teilnehmen, diese aufzeichnen, transkribieren und analysieren. Besonderer Fokus liegt auf der Extraktion von Aktionspunkten, Entscheidungen und wichtigen Themen aus dem Gespräch.

Link: <https://fireflies.ai/>

4. GitHub Copilot

Ein "KI-Paarprogrammierer" für Softwareentwickler. Direkt im Code-Editor (z.B. VS Code) schlägt Copilot basierend auf dem Kontext (Kommentare, umgebender Code) ganze Codezeilen oder Funktionen vor, was die Entwicklungsgeschwindigkeit erheblich steigern kann.

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

5. Replit Ghostwriter

Ein integrierter KI-Coding-Assistent innerhalb der cloudbasierten Entwicklungsumgebung Replit. Ghostwriter bietet Code-Vervollständigung, Code-Erklärung, Debugging-Hilfe und kann Code basierend auf Anweisungen in natürlicher Sprache generieren.

Link: <https://replit.com/ghostwriter>

6. Microsoft Copilot für Microsoft 365

Eine tiefgreifende KI-Integration in die Microsoft Office Suite (Word, Excel, PowerPoint, Outlook, Teams). Copilot kann Dokumente entwerfen, Daten in Excel analysieren, Präsentationen erstellen, E-Mails formulieren und Meetings

zusammenfassen, basierend auf den Daten des Unternehmens.

Link: <https://www.microsoft.com/de-de/microsoft-365/copilot-for-microsoft-365>

7. Zoom AI Companion

Bietet KI-gestützte Funktionen direkt innerhalb von Zoom-Meetings und -Chats. Dazu gehören automatische Meeting-Zusammenfassungen, die Identifizierung von Aktionspunkten, Hilfe beim Verfassen von Chat-Nachrichten und zukünftig auch Echtzeit-Übersetzungen.

Link: <https://zoom.us/ai-companion>

8. Google Workspace Duet AI (jetzt Teil von Gemini)

Googles Pendant zu Microsoft Copilot für die eigene Office-Suite (Docs, Sheets, Slides, Meet). Die KI (jetzt unter dem Gemini-Banner) hilft beim Schreiben in Docs, Erstellen von Tabellen in Sheets, Generieren von Bildern für Slides und beim Zusammenfassen von Meetings.

<https://workspace.google.com/solutions/ai/>

9. KI-Kalenderassistenten (z.B. Clockwise, Reclaim.ai)

Diese Tools analysieren den Kalender und die Arbeitsmuster des Nutzers, um automatisch Fokuszeiten zu blocken, Meetings intelligent zu verschieben (innerhalb vorgegebener Grenzen) und Konflikte aufzulösen. Ziel ist eine bessere Zeitplanung und mehr ungestörte Arbeitszeit.

<https://www.getclockwise.com/>,

<https://reclaim.ai/>

10. QuillBot

Ein KI-Schreibwerkzeug mit Fokus auf Paraphrasierung. Es hilft dabei, Sätze oder ganze Absätze umzuformulieren, um Plagiate zu vermeiden, die Klarheit zu verbessern oder den Stil anzupassen. Bietet auch Funktionen zum Zusammenfassen und zur Grammatikprüfung.

Link: <https://quillbot.com/>

11. Texterkennung (OCR) in Apps (z.B. Google Lens, Adobe Scan)

Optical Character Recognition (OCR) nutzt KI, um Text in Bildern (z.B. Fotos von Dokumenten, Schildern) zu erkennen. Diese Technologie ermöglicht es, Text aus Bildern zu kopieren, zu übersetzen oder durchsuchbar zu machen, was z.B. das Digitalisieren von Papierdokumenten enorm erleichtert.

Links: <https://lens.google/>,

<https://acrobat.adobe.com/de/de/mobile/scanner-app.html>

Kreativität & Content

1. Midjourney

Ein sehr populärer KI-Bildgenerator, der für seine oft hochästhetischen und künstlerisch anspruchsvollen Ergebnisse bekannt ist. Nutzer interagieren mit Midjourney über Textbefehle (Prompts) auf der Chat-Plattform Discord, um einzigartige Bilder zu erschaffen.

Link: <https://www.midjourney.com/> (Zugang über Discord)

2. DALL-E 3 (OpenAI)

Ein weiteres führendes Modell zur Erzeugung von Bildern aus Textbeschreibungen, entwickelt von OpenAI. DALL-E 3 ist bekannt dafür, Prompts oft sehr genau umzusetzen und auch Text in Bildern darstellen zu können. Es ist u.a. in ChatGPT Plus und Microsofts Image Creator integriert.

Link: <https://openai.com/dall-e-3/>

3. Stable Diffusion (z.B. DreamStudio)

Ein einflussreiches Open-Source-Modell zur Bildgenerierung. Die Offenheit ermöglicht eine breite Nutzung und Anpassung durch die Community. DreamStudio ist eine benutzerfreundliche Weboberfläche, um Stable Diffusion zu nutzen, es gibt aber viele weitere Implementierungen.

Link: <https://beta.dreamstudio.ai/>

4. Canva Magic Studio

Canva, eine Design-Plattform, hat eine Reihe von KI-Funktionen unter dem Namen "Magic Studio" integriert. Dazu gehören Texterstellung (Magic Write), Bildgenerierung, Objektentfernung/-ersetzung in Fotos, automatische Designvorschläge und sogar Text-zu-Video-Funktionen.

Link: <https://www.canva.com/magic-studio/>

5. Adobe Firefly

Adobes eigene generative KI-Engine, die speziell für kreative Arbeitsabläufe entwickelt wurde und auf ethisch unbedenklichen Daten trainiert sein soll. Firefly ist in Photoshop ("Generative Fill"), Illustrator und Adobe Express integriert und bietet Funktionen zur Bilderstellung und -bearbeitung per Textprompt.

Link: <https://firefly.adobe.com/>

6. RunwayML

Eine Online-Plattform, die sich auf KI-gestützte Videobearbeitung spezialisiert hat. Runway bietet Werkzeuge wie Text-zu-Video-Generierung, automatische Objektentfernung aus Videos (Inpainting), Zeitlupeneffekte (Frame Interpolation) und vieles mehr.

Link: <https://runwayml.com/>

7. Synthesia

Ermöglicht die Erstellung von Videos mit realistisch wirkenden KI-Avataren, die einen vorgegebenen Text sprechen. Dies ist nützlich für Schulungsvideos, Produktpräsentationen oder Erklärvideos, die schnell und in vielen Sprachen produziert werden können, ohne echte Schauspieler filmen zu müssen.

Link: <https://www.synthesia.io/>

8. Lensa AI

Eine mobile Bildbearbeitungs-App, die besonders durch ihre Funktion "Magic Avatars" bekannt wurde. Nutzer laden eigene Fotos

hoch, und die KI generiert daraus eine Reihe von künstlerisch stilisierten Porträts und Avataren in verschiedenen Stilen.

Link: <https://prisma-ai.com/lensa>

9. Soundraw / Amper Music

KI-Musikgeneratoren, die es Nutzern ermöglichen, lizenzfreie Musikstücke zu erstellen, indem sie Parameter wie Stimmung, Genre, Instrumente und Länge vorgeben. Nützlich für Content Creator, die Hintergrundmusik für Videos oder Podcasts benötigen. (Amper wurde von Shutterstock übernommen und integriert).

Link: <https://soundraw.io/>

10. ElevenLabs

Ein führender Anbieter für KI-basierte Text-to-Speech (TTS) Technologie. ElevenLabs kann extrem realistische und emotional nuancierte menschliche Stimmen aus geschriebenem Text generieren. Es bietet auch die Möglichkeit, eigene Stimmen zu klonen (Voice Cloning).

Link: <https://elevenlabs.io/>

11. Descript

Ein innovatives Audio- und Videobearbeitungstool, das wie ein Textdokument funktioniert. Es transkribiert die Aufnahme automatisch, und Bearbeitungen (Schneiden, Kopieren, Einfügen) erfolgen direkt im Text. KI-Funktionen wie "Overdub" (Stimme korrigieren/ergänzen) und "Studio Sound" (Audioqualität verbessern) sind integriert.

Link: <https://www.descript.com/>

12. Foto-Editoren (z.B. Luminar Neo, Photoshop AI)

Moderne Bildbearbeitungsprogramme integrieren immer mehr KI-Funktionen. Beispiele sind das automatische Austauschen des Himmels, das intelligente Entfernen oder Hinzufügen von Objekten (Generative Fill in Photoshop), die Verbesserung von Porträts oder die automatische Anpassung von Belichtung und Farben.

Links: <https://skylum.com/luminar>, <https://www.adobe.com/products/photoshop.html>

Unterhaltung & Shopping

1. Netflix Empfehlungen

Das Herzstück von Netflix ist sein Empfehlungssystem. KI analysiert detailliert Ihr Sehverhalten (angeschaute Titel, Bewertungen, Pausen, Suchanfragen) sowie das Verhalten ähnlicher Nutzer, um Ihnen möglichst passende Filme und Serien vorzuschlagen und Sie auf der Plattform zu halten.

Link: <https://www.netflix.com/>

2. Spotify Empfehlungen & Playlists

Spotify nutzt KI (insbesondere "Collaborative Filtering" und Audioanalyse), um Ihren Musikgeschmack zu verstehen. Darauf basierend erstellt es personalisierte Playlists wie "Discover Weekly" oder "Release Radar" und schlägt Ihnen neue Songs und Künstler vor, die Ihnen gefallen könnten. Link: <https://www.spotify.com/>

3. YouTube Empfehlungen

Ähnlich wie Netflix und Spotify verwendet YouTube hochentwickelte KI-Algorithmen, um Videos für Ihre Startseite und die "Nächstes Video"-Vorschläge auszuwählen. Die KI berücksichtigt Ihren Wiedergabeverlauf, Likes, Abonnements und Trends, um die Verweildauer zu maximieren.

Link: <https://www.youtube.com/>

4. Amazon Produktempfehlungen

Wenn Sie auf Amazon stöbern, analysiert eine KI Ihr Klick- und Kaufverhalten, sowie das von Millionen anderer Kunden. Daraus generiert sie personalisierte Produktvorschläge ("Kunden, die diesen Artikel angesehen haben...", "Wird oft zusammen gekauft...") um den Umsatz zu steigern. Link: <https://www.amazon.de/>

5. E-Commerce Chatbots

Viele Online-Shops setzen KI-gesteuerte Chatbots ein, um häufig gestellte Fragen (FAQs) zu Versand, Rückgabe oder Produktinformationen rund um die Uhr zu beantworten. Sie können auch bei der Produktsuche helfen oder einfache Bestellprozesse begleiten.

(Auf vielen Shop-Websites zu finden)

6. Personalisierte Werbung

Ein Großteil der Online-Werbung wird mittels KI personalisiert. Algorithmen analysieren Ihr Surfverhalten, Suchanfragen, Standortdaten und demografische Merkmale, um Ihnen möglichst relevante Anzeigen zu präsentieren (Targeting und Retargeting).

(Überall im Web)

7. Dynamische Preisgestaltung (z.B. Fluggesellschaften [diverse Logos], Uber)

KI-Systeme passen Preise für Produkte oder Dienstleistungen in Echtzeit an. Faktoren wie Nachfrage, Angebot, Tageszeit, Wettbewerbspreise und sogar das individuelle Kundenprofil können einfließen, um den Ertrag zu optimieren (z.B. bei Flugtickets oder Fahrdiensten).

(Websites von Airlines, Uber App)

8. KI in Videospielen

Künstliche Intelligenz ist entscheidend für moderne Videospiele. Sie steuert das Verhalten von Nicht-Spieler-Charakteren (NPCs), passt den Schwierigkeitsgrad dynamisch an das Können des Spielers an, generiert prozedural Spielwelten oder verbessert die Grafik (z.B. DLSS von Nvidia). (In vielen modernen Spielen integriert)

Alltag & Zuhause

1. Google Maps / Waze / Apple Karten

Navigations-Apps nutzen KI intensiv. Sie analysieren Echtzeit-Verkehrsdaten (von anderen Nutzern), historische Daten und Ereignismeldungen, um die schnellste Route zu berechnen, Staus zu umfahren und präzise Ankunftszeiten vorherzusagen.

Links: <https://www.google.com/maps>,
<https://www.waze.com/>,
<https://www.apple.com/de/maps/>

2. Smarte Thermostate (z.B. Nest, Tado)

Diese Geräte lernen Ihre Heiz- und Kühlgewohnheiten sowie Ihre Anwesenheitszeiten (oft über Geofencing per Smartphone). Die KI erstellt dann automatisch Heizpläne, um Komfort zu gewährleisten und gleichzeitig Energie zu sparen, indem unnötiges Heizen vermieden wird.

Links: https://store.google.com/product/nest_learning_thermostat_3rd_gen,
<https://www.tado.com/de-de/>

3. Roboterstaubsauger (z.B. Roomba, Roborock)

Moderne Saugroboter nutzen KI (oft kombiniert mit Sensoren wie LiDAR oder Kameras) zur intelligenten Navigation. Sie erstellen Karten Ihrer Wohnung, erkennen Hindernisse, planen effiziente Reinigungsrouten und lernen, welche Bereiche besonders schmutzig sind.

(Hersteller-Websites)

4. Fitness Tracker & Smartwatches (z.B. Fitbit, Apple Watch, Garmin)

Diese Wearables sammeln kontinuierlich Daten über Ihre Aktivität, Herzfrequenz, Schlafqualität etc. KI-Algorithmen analysieren diese Daten, um Muster zu erkennen, per-

sonalisierte Einblicke in Ihre Gesundheit zu geben, Trainingsfortschritte zu verfolgen und potenzielle Unregelmäßigkeiten zu melden. (Hersteller-Websites)

5. Gesichtserkennung zum Entsperren (z.B. Face ID, Android Face Unlock)

Smartphones nutzen KI-Modelle (Neuronale Netze), um Ihr Gesicht anhand vieler einzigartiger Merkmale zu erkennen und das Gerät sicher zu entsperren. Diese Systeme lernen und werden auch bei leichten Veränderungen (Brille, Bart) zuverlässiger. (Integrierte Handy-Funktion)

6. Spracherkennung für Diktate

Betriebssysteme (Windows, macOS, iOS, Android) und viele Apps enthalten eine Diktierfunktion, die gesprochene Sprache mittels KI (Automatic Speech Recognition - ASR) in geschriebenen Text umwandelt. Die Genauigkeit hat sich in den letzten Jahren enorm verbessert.

(Betriebssystem-Funktion)

Sicherheit & Finanzen

1. Kreditkartenbetrugserkennung

Banken und Kreditkarten-Unternehmen setzen KI-Systeme ein, die Transaktionen in Echtzeit analysieren. Sie lernen typische Ausgabemuster und erkennen Abweichungen (z.B. ungewöhnlicher Ort, hoher Betrag), um potenziellen Betrug zu identifizieren u. zu blockieren. (Hintergrundfunktion bei Banken/Kreditkartenanbietern)

2. Antivirus-Software (Verhaltensanalyse)

Neben signaturbasierten Scans nutzen moderne Antivirenprogramme KI zur Verhaltensanalyse. Sie überwachen Programme auf verdächtige Aktionen (z.B. Verschlüsselung vieler Dateien) und können so auch neue, noch unbekannte Malware (Zero-Day-Exploits) erkennen. (z.B. Norton, Bitdefender, Kaspersky)

3. CAPTCHA & reCAPTCHA

Diese "Mensch-oder-Maschine"-Tests nutzen KI, um automatisierte Bots von menschlichen Nutzern zu unterscheiden. Googles reCAPTCHA analysiert im Hintergrund Nutzerverhalten (Mausbewegungen, Klickmuster), um oft ohne sichtbare Aufgabe entscheiden zu können. Link: <https://www.google.com/recaptcha/>

4. Robo-Advisors (z.B. Scalable Capital, Betterment)

Online-Vermögensverwaltungen, die KI und Algorithmen nutzen, um Anlagestrategien zu entwickeln und Portfolios automatisch zu verwalten. Basierend auf Risikobereitschaft und Zielen des Kunden wird ein diversifiziertes Portfolio erstellt und regelmäßig angepasst (Rebalancing).

Links: <https://de.scalable.capital/>,
<https://www.betterment.com/>

5. Banking App Sicherheitsfeatures

Viele Banking-Apps integrieren KI-basierte Sicherheitsmechanismen. Diese können ungewöhnliche Login-Versuche (z.B. von neuem Gerät oder Standort), verdächtige Transaktionsmuster oder potenzielle Phishing-Versuche erkennen und den Nutzer warnen oder zusätzliche Verifizierungsschritte einleiten.

(Integrierte Funktion vieler Banking-Apps)

Lernen & Entwicklung

1. Duolingo

Die populäre Sprachlern-App nutzt KI, um den Lernprozess zu personalisieren. Algorithmen analysieren die Antworten des Nutzers, um den Schwierigkeitsgrad anzupassen und zu entscheiden, welche Vokabeln oder Grammatikregeln wann wiederholt werden sollten (Spaced Repetition).

Link: <https://de.duolingo.com/>

2. Khan Academy

Diese gemeinnützige Lernplattform setzt KI ein, um personalisierte Lernpfade für Schüler in verschiedenen Fächern (insbesondere Mathematik) zu erstellen. Das System passt Übungen an das individuelle Tempo und Verständnis des Lernenden an und gibt gezielte Hilfestellungen.

Link: <https://de.khanacademy.org/>

3. Perplexity AI

Eine KI-gestützte Suchmaschine, die darauf ausgelegt ist, direkte, informative Antworten auf Fragen zu geben, inklusive Quellenangaben. Sie eignet sich gut für Recherche, da sie Informationen aus dem Web zusammenfasst und übersichtlich darstellt.

Link: <https://www.perplexity.ai/>

4. Phind

Eine spezialisierte KI-Suchmaschine für Programmierer und technische Fragen. Phind durchsucht das Web nach relevanten Informationen und Code-Beispielen und generiert präzise Antworten und Erklärungen, oft schneller als eine traditionelle Suchmaschine.

Link: <https://www.phind.com/>

5. WolframAlpha

Weniger eine Suchmaschine als eine "Computational Knowledge Engine". WolframAlpha nutzt KI und eine riesige kuratierte Wissensdatenbank, um Fragen durch Berechnung zu beantworten. Es kann mathematische Probleme lösen, Daten analysieren, Fakten vergleichen und komplexe Berechnungen durchführen.

Link: <https://www.wolframalpha.com/> 

Die Selbstverbesserung der Künstlichen Intelligenz

KI-Selbstverbesserung ist ein fortgeschrittener Bereich der KI, der sich mit der Entwicklung von Systemen beschäftigt, die fähig sind, **ihre eigenen Leistungsfähigkeit und Effizienz zu optimieren**, ohne direkte menschliche Eingriffe. Hier einige Aspekte dieser Neuerscheinung:

Die Entwicklung von KI-Systemen, die in der Lage sind, sich selbst zu verbessern, wird von einer Vielzahl von Forschern, Wissenschaftlern und Entwicklern in der Künstlichen Intelligenz vorangetrieben.

Die Entwicklung von KI-Systemen, die sich selbst verbessern können, ist ein **multidisziplinäres Unterfangen**, das Experten aus verschiedenen Bereichen – Informatik, Neurowissenschaften, Mathematik und Philosophie – zusammenbringt.

Es ist ein dynamisches und schnell wachsendes Feld, das kontinuierliche Innovation und Zusammenarbeit erfordert.

Diese Bemühungen sind nicht auf einzelne Unternehmen oder Organisationen beschränkt, sondern umfassen eine breite Palette von Akteuren aus Akademie und Industrie:

Akademische Institutionen:

Universitäten und Forschungsinstitute weltweit spielen eine entscheidende Rolle bei der Grundlagenforschung im Bereich der selbstverbessernden KI. Institutionen wie das MIT (Massachusetts Institute of Technology), Stanford University, und das Google Brain Team (das eng mit der Stan-

ford University verbunden ist) führen bahnbrechende Forschungen durch.

Konferenzen und Workshops wie NIPS (Neural Information Processing Systems) und ICLR (International Conference on Learning Representations) bieten Plattformen für Forscher, um ihre neuesten Erkenntnisse und Fortschritte in der selbstverbessernden KI zu teilen.

Technologieunternehmen:

Große Tech-Unternehmen wie **Google, Microsoft, Facebook (Meta), IBM** und **Amazon** investieren erheblich in die Forschung und Entwicklung von selbstverbessernder KI, oft durch spezialisierte Teams oder Labore wie Google's DeepMind.

Start-ups und kleinere Unternehmen konzentrieren sich ebenfalls auf innovative Ansätze in diesem Bereich, oft mit dem Ziel, spezifische Probleme zu lösen oder neue Technologien zu entwickeln.

Regierungen und Verteidigungsbehörden:

Regierungsorganisationen und Verteidigungsbehörden wie die DARPA (Defense Advanced Research Projects Agency) in den USA fördern Forschung und Entwicklung in selbstverbessernder KI, insbesondere mit Blick auf militärische und sicherheitsrelevante Anwendungen.

Zusammenarbeit und Offene Projekte:

Offene Quellcode-Projekte und Kooperationen zwischen verschiedenen Forschungseinrichtungen und Unternehmen fördern den Austausch von Wissen und beschleunigen den Fortschritt in diesem Feld.

Einige Mechanismen der KI-Selbstverbesserung

Die Zukunft der KI wird maßgeblich von der Fähigkeit zur Selbstverbesserung geprägt sein. BStaR ist ein wichtiger Schritt in diese Richtung und zeigt, welches Potenzial in dieser Technologie steckt.

Es bleibt spannend zu beobachten, wie sich diese Entwicklung weiter fortsetzt und welche Auswirkungen sie auf unsere Welt haben wird.

KI-Selbstverbesserung hat das Potenzial, die Effizienz und Effektivität von KI-Systemen erheblich zu steigern, erfordert jedoch auch eine sorgfältige Abwägung von Risiken, um sicherzustellen, dass diese Technologie verantwortungsvoll und sicher eingesetzt wird. Einige Mechanismen der KI-Selbstverbesserung sind folgende:

1. Algorithmische Anpassung:

Ein selbstverbesserndes KI-System kann seine internen Algorithmen ändern, um bessere Entscheidungen zu treffen oder effizienter zu arbeiten.

Ein System, das Muster in Daten erkennt, kann seine Mustererkennungsalgorithmen anpassen, um feinere oder komplexere Muster zu erkennen.

2. Lernen durch Meta-Lernen:

Dieses Konzept, auch bekannt als "Lernen zu lernen", ermöglicht es einer KI, Strategien zu entwickeln, wie sie am besten lernen und sich an neue Aufgaben anpassen kann.

Ein Meta-Lernsystem könnte beispielsweise lernen, welche Lernstrategien für bestimmte Arten von Aufgaben am effektivsten sind.

3. Strukturanpassung:

Fortgeschrittene Formen der KI-Selbstverbesserung erlauben es Systemen, nicht nur ihre Algorithmen, sondern auch ihre interne Struktur zu verändern.

Dies könnte die Hinzufügung neuer Neuronen in Neuronalen Netzwerken oder die Modifikation bestehender Verbindungen umfassen, um die Leistung zu steigern.

Herausforderungen und Überlegungen:

Sicherheit und Kontrolle:

Eine der größten Herausforderungen bei selbstverbessernder KI ist die Sicherstellung, dass das System innerhalb sicherer Grenzen operiert und seine Verbesserungen nicht unerwünschte oder gefährliche Auswirkungen haben...

Transparenz und Verständlichkeit:

Selbstverbessernde Systeme können komplex und schwer zu verstehen sein, was Fragen zur Transparenz und Erklärbarkeit ihrer Entscheidungen aufwirft.

Ethik und Verantwortung:

Es gibt bedeutende ethische Überlegungen hinsichtlich der Entwicklung und Verwendung von KI-Systemen, die in der Lage sind, sich selbst zu verbessern, einschließlich der Verantwortung für ihre Handlungen. 🐼

B-STaR: Grenzenloses Lernen und die Definition der KI-Selbstverbesserung

Neuere Ansatz in der Entwicklung von Künstlicher Intelligenz: Die Rolle der KI-Selbstverbesserung mit dem **Balanced Self-Taught Reasoner**

Die Entwicklung Künstlicher Intelligenz schreitet rasant voran. Immer schneller werden neue Meilensteine erreicht, die die Grenzen des Möglichen verschieben. Ein besonders faszinierendes und zugleich herausforderndes Gebiet ist die **Selbstverbesserung von KI**.

Hier setzt **B-STaR**, eine neue KI-Architektur, an und stellt bisherige Annahmen über die Möglichkeiten des **Maschinellen Lernens** in Frage.

B-STaR zeichnet sich durch einen neuartigen Ansatz aus, der es der KI ermöglicht, ihre eigenen Algorithmen und Strukturen kontinuierlich zu optimieren.

Im Gegensatz zu herkömmlichen KI-Systemen, die auf vorgegebenen Datensätzen trainiert werden und deren Lernfähigkeit begrenzt ist, kann **B-STaR** eigenständig neue Lernstrategien entwickeln und anwenden.

Dieser iterative Prozess der Selbstverbesserung führt zu einer exponentiellen Steigerung der Leistungsfähigkeit und eröffnet ungeahnte Möglichkeiten.

Die Grundlage von **B-STaR** bildet ein komplexes Netzwerk aus miteinander verbundenen Modulen, die ähnlich wie Neuronen im menschlichen Gehirn funktionieren. Diese Module

B-STaR steht für "**Balanced Self-Taught Reasoner**". **B-STaR** (oft auch "**B-STAR**" geschrieben) ist ein spezifisches Forschungsprojekt im Bereich der KI. Es befasst sich mit der zentralen Herausforderung bei **selbstlernenden KI-Systemen**, insbesondere großen Sprachmodellen (Large Language Models, LLMs), die komplexe Denkaufgaben lösen sollen (wie z.B. mathematische Probleme oder Programmierung). Dabei geht es um das Balancieren von "Exploration" und "Exploitation":

- **Exploration:** Die Fähigkeit des Modells, vielfältige neue Lösungsansätze oder "Gedankengänge" zu generieren.
- **Exploitation:** Die Fähigkeit des Modells, anhand von Belohnungen (z.B. ob die generierte Antwort korrekt ist) die vielversprechendsten Ansätze zu erkennen und zu nutzen.

Die Forschung hat gezeigt, dass bei rein selbstlernenden Methoden die Explorationsfähigkeit von Modellen schnell nachlassen kann und die Effektivität der Nutzung von Belohnungen abnimmt.

B-STaR begegnet diesem Problem, indem es die Konfigurationen des KI-Modells während des Trainings dynamisch anpasst. Es verwendet Metriken, um den aktuellen Zustand von **Exploration** und **Exploitation** zu überwachen und die Parameter so zu steuern, dass ein besseres Gleichgewicht erreicht wird.

Ziel von **B-STaR** ist es, eine effektivere Selbstverbesserung von KI-Modellen zu ermöglichen, die über mehr Iterationen hinweg konsistente Leistungssteigerungen erzielt, im Gegensatz zu früheren Methoden, die oft nach wenigen Trainingszyklen stagnieren.

B-STaR ist also ein Forschungsansatz zur Verbesserung der Lernfähigkeit und des Denkvermögens von KI-Modellen durch dynamisches Management von Exploration und Exploitation im Selbstlernprozess.

sind in der Lage, Informationen zu verarbeiten, zu speichern und miteinander auszutauschen. Durch einen ausgeklügelten Feedback-Mechanismus lernt B-STaR, welche Module und Verbindungen für die Lösung bestimmter Aufgaben am effektivsten sind und passt seine Architektur entsprechend an.

Diese Fähigkeit zur Selbstverbesserung wirft jedoch auch ethische und sicherheitsrelevante Fragen auf. Wie kann sichergestellt werden, dass eine KI, die sich selbstständig weiterentwickelt, die Kontrolle behält und im Sinne des Menschen handelt?

Die Entwickler von B-STaR sind sich dieser Herausforderungen bewusst und arbeiten an Sicherheitsmechanismen, die ein unkontrolliertes Wachstum der KI verhindern sollen. Dazu gehören unter anderem die Implementierung von ethischen Richtlinien und die ständige Überwachung des Lernprozesses.

B-STaR hat das Potenzial, die Welt grundlegend zu verändern. Von der Medizin über die Forschung bis hin zur Industrie – die Anwendungsmöglichkeiten sind nahezu unbegrenzt. Die Fähigkeit zur Selbstverbesserung eröffnet neue Wege zur Lösung komplexer Probleme und zur Entwicklung innovativer Technologien. Gleichzeitig ist es wichtig, die damit verbundenen Risiken ernst zu nehmen und verantwortungsvoll mit dieser neuen Technologie umzugehen.

Die Zukunft der KI wird maßgeblich von der Fähigkeit zur Selbstverbesserung geprägt sein. B-STaR ist ein wichtiger Schritt in diese Richtung und zeigt, welches Potenzial in dieser Technologie steckt. Es bleibt spannend zu beobachten, wie sich diese Entwicklung weiter fortsetzt und welche Auswirkungen sie auf unsere Welt haben wird.

Wer hat B-STaR erfunden?

Der Artikel *B-STaR: Monitoring and Balancing Exploration and Exploitation in Self-Taught Reasoners* wurde am 23. Dezember 2024 veröffentlicht. Das B-STaR Framework wurde von einer Gruppe von Forschenden entwickelt. Die Autoren sind:

- Weihao Zeng
- Lulu Zhao
- Yijun Wang

und drei weitere Ko-Autoren die in der Publikation zu finden sind. Es handelt sich also um eine Kooperation von mehreren Forschenden, die dieses spezifische KI-Framework konzipiert haben.

Die Zugehörigkeit der Autoren ist typischerweise in der wissenschaftlichen Publikation angegeben (oft einer Universität oder Forschungseinrichtung), aber diese spezifischen Details waren nicht vollständig enthalten. Die Veröffentlichung des Papers auf Plattformen wie arXiv und bei Konferenzen wie ICLR (International Conference on Learning Representations) ist der standardmäßige Weg, neue Forschungsergebnisse im Bereich der KI zu präsentieren.

Weihao Zeng ist einer der Hauptautoren des Forschungsartikels, der das B-STaR-Framework vorstellt: “B-STaR: Monitoring and Balancing Exploration and Exploitation in Self-Taught Reasoners”. Weihao Zeng ist ein KI-Forscher, insbesondere in großen Sprachmodellen (LLMs) und deren Selbstverbesserung sowie im Bereich des Task-Oriented Dialogue Systems. Seine Affiliationen umfassen unter anderem:

- Beijing University of Posts and Telecommunications (BUPT), wo er als Student (Master und Bachelor) geführt wird.
- Microsoft, wo er als Praktikant (Intern) angegeben ist (Zeitraum 2023-2025).

Es scheint, dass Weihao Zeng zum Zeitpunkt der Entwicklung und Veröffentlichung des B-STaR-Papiers an der Beijing University of Posts and Telecommunications forschte und gleichzeitig bei Microsoft als Praktikant tätig war.

Vorteile und Nachteile der Selbstmodifikation bei der Künstlichen Intelligenz

Die Selbstmodifikation bei Künstlicher Intelligenz bringt sowohl Gewinne als auch Risiken. Auf der positiven Seite könnte eine hochintelligente KI in der Lage sein, ihre eigenen Fähigkeiten zu verbessern, effektiver Probleme zu lösen und möglicherweise sogar die menschliche Kreativität zu übertreffen. Dies könnte zu Durchbrüchen in Wissenschaft, Medizin und Technologie führen.

Auf der anderen Seite besteht das Risiko, dass eine solche KI außer Kontrolle gerät und

ihre Fähigkeiten auf eine Weise einsetzt, die für Menschen schädlich oder gefährlich sind. Die Sicherheit und ethische Verantwortung bei der Entwicklung und Nutzung von KI-Systemen sind daher von entscheidender Bedeutung, um etwaige Risiken und negative Auswirkungen zu minimieren.

Jetzt wissen wir, wie sich Maschinelle Intelligenz selbst modifizieren kann und welche Auswirkungen eine mögliche "Intelligenzexplosion" haben könnte. Lassen Sie uns nun einen Blick auf die Intelligenzexplosion selbst werfen und die Herausforderungen und Chancen, die sie mit sich bringt. 🤖



In der nächsten Ausgabe dieser Newsletter – Nr. 13 – Informationen über Roboter...

Die verblüffende Ambitionen des Startups *Mechanize*

Ein revolutionärer Durchbruch oder Silicon Valley's dreisteste Satire? Der kuriose Fall vom kühnen KI-Unternehmer **Tamay Besiroglu**, der **Mechanize** neulich gründete...

Das Silicon Valley hat die Geburt von vielen Startups erlebt, dessen Innovationen vielversprechend waren, um die Zukunft umzugestalten; aber gelegentlich taucht eine Prämisse auf, die so provokativ ist, dass es schwierig ist, zu unterscheiden, ob es sich um echtem Ehrgeiz oder nur um Satire handelt.

Da ist der kuriose Fall vom kühnen neuen Unternehmen *Mechanize*, welches vom gefeierten KI-Forscher **Tamay Besiroglu** gegründet wurde.

Tamay Besiroglu, bekannt für die Schaffung der einflussreichen gemeinnützigen KI-Forschungsorganisation *Epoch*, hat kürzlich *Mechanize* vorgestellt, ein Unternehmen mit dem außergewöhnlichen Ziel, „alle menschlichen Arbeiten“ zu automatisieren...

.....

Er stellt sich sogar *die vollständige Automatisierung der globalen Wirtschaft* vor. Laut **Tamay Besiroglu** beabsichtigt das Startup, die notwendige Infrastruktur bereitzustellen – Daten, ausgefeilte KI-Auswertungen und realistische virtuelle Umgebungen –, damit Roboter und Digital-Agenten menschliche Arbeiter in fast allen erdenklichen beruflichen Bereichen **vollständig ersetzen können**.

Bei der Enthüllung des enormen Ausmaßes der Bestrebungen von *Mechanize* wies **Tamay Besiroglu** auf das Ausmaß seines potenziellen Marktes hin. Er berechnete die globalen Jahreslöhne aller Arbeiter auf etwa **60 Billionen Dol-**



Der heute berühmte KI-Forscher **Tamay Besiroglu** (oben) hat mit der Gründung von *Mechanize*, einem Startup mit der umstrittenen Mission, „die vollständige Automatisierung aller Arbeit“ und „die vollständige Automatisierung der Wirtschaft“ zu erreichen, viele Kontroversen ausgelöst.

Das Unternehmen erntet Kritik sowohl für seine Ziele als auch für den möglichen Schaden am Ruf von *Epoch AI*, dem angesehenen Forschungsinstitut, das **Tamay Besiroglu** mitbegründet hat.

Tamay Besiroglu, Mitbegründer von *Mechanize* zusammen mit Matthew Barnett und Ege Erdil, stellt sich eine Zukunft vor, in der KI-gesteuerte Automatisierung „großen Überfluss“ und „viel höhere Lebensstandards“ schafft.

lar, wobei allein die amerikanischen Arbeiter etwa **18 Billionen Dollar** dieser Gesamtsumme ausmachen.

Trotz des expansiven langfristigen Ziels stellte **Tamay Besiroglu** den Reportern jedoch klar, dass der unmittelbare Fokus seines Unternehmens eher auf Angestellten- und nicht auf Handarbeit liegt, was von Natur aus komplexere Roboterlösungen erfordern würde.

Sein wirtschaftliches Argument konzentriert sich darauf, Arbeit in eine akkumulierbare Ressource zu verwandeln – im Gegensatz zur traditionellen Wirtschaft, in der Arbeit fixiert ist, ermöglicht KI die Schaffung digitaler Arbeitskräfte durch Investitionen, was potenziell ein explosives Wirtschaftswachstum auslösen könnte.

.....
Sofortige und weit verbreitete Kritik an dem Konzept brach aus, wobei viele in der Technologie-Community von *Mechanizes* erklärter Mission beunruhigt waren.

Prominente Stimmen drückten ihre Besorgnis darüber aus, dass das Projekt genau die Hypererweiterung von Automatisierungsängsten darstellt, die Experten seit dem Aufstieg der KI alarmiert haben.

Anthony Aguirre war unter denen, die Enttäuschungen ausdrückten, und erklärte öffentlich, dass er zwar die früheren Beiträge des Gründers respektiert, das Unternehmen aber eher einen tiefgreifenden Verlust als Fortschritt darstellen könnte, wobei er betont, dass die Automatisierung in diesem Ausmaß die Menschheit im Allgemeinen negativ beeinflussen könnte, anstatt sie zu bereichern...

Erschwerend kommt eine Kontroverse hinzu, welche die enge Verbindung zwischen *Mechanize* und *Epoch* aufgedeckt hat, die zuvor als unparteiische Organisation angesehen wurde, die sich weitgehend der Bewertung von KI-Technologien kritisch und transparent widmet.

Die Stimmen der Kritiker

Kritiker, darunter prominente Branchenvertreter wie **Oliver Habryka**, äußerten prompt Bedenken über mögliche Interessenkonflikte und stellten die Neutralität von *Epoch* in Frage, da sein Gründer (**Besiroglu**) nun direkt von der Verbesserung der KI-Fähigkeiten durch *Mechanize* profitieren würde.

In der Tat ist dies nicht der erste Fall von Kontroversen um *Epoch*, die zuvor auf die Lupe genommen wurde, weil sie zunächst ihre Unterstützung von **OpenAI** bei einem Großprojekt heruntergespielt hatte.

Trotz der Kritik hat *Mechanize* eine bedeutende erste Unterstützung von einer beeindruckenden Liste von Investoren erhalten, darunter **Nat Friedman**, **Daniel Gross**, **Patrick Collison**, **Dwarkesh Patel**, **Jeff Dean**, **Sholto Douglas** und **Marcus Abramovitch**, ein Partner beim Krypto-Hedgefonds **AltX** und selbsternannter effektiver Altruist.

.....
Marcus Abramovitch, der sein Engagement bestätigte, lobte die außergewöhnlichen Fähigkeiten des Gründungsteams und deutete an, dass sie die Auswirkungen von KI gründlicher durchdacht hatten als jeder, den er auf dem Gebiet getroffen hat.

Als Reaktion auf die Gegenreaktion argumentiert **Tamay Besiroglu** ernsthaft, dass die Vision seines Start-ups, anstatt eine Bedrohung für die menschlichen Lebensgrundlagen darzustellen, das gesellschaftliche Wohlergehen tatsächlich durch explosives Wirtschaftswachstum, deutlich verbesserte Lebensstandard und die Schaffung innovativer Produkte und Dienstleistungen, die derzeit unvorstellbar sind, ernsthaft erhöhen kann.

Tamay Besiroglu zitiert veröffentlichte Forschungsergebnisse, die behaupten, dass eine umfassende Automatisierung einen beispiellosen globalen Überfluss freisetzen könnte.

Unterstützung von Größen der KI-Technologie

Trotz Kritik hat **Tamay Besiroglu** beeindruckende Unterstützung von Tech-Größen wie **Nat Friedman**, **Daniel Gross**, **Patrick Collison**, **Dwarkesh Patel**, **Jeff Dean** und anderen gewonnen.

Trotzdem besteht große Skepsis darüber, ob *Mechanize* tatsächlich einen ernsthaften Versuch darstellt, die Wirtschaft zu revolutionieren, oder lediglich die Entwicklung von KI-Fähigkeiten unter einem anderen Banner beschleunigt.

Dennoch sind die Skeptiker nach wie vor nicht überzeugt, da sie beobachten, dass eine vollständig automatisierte Belegschaft die Kaufkraft von Arbeitslosen drastisch reduzieren könnte, was letztlich die Verbrauchernachfrage und die breitere Wirtschaft schwächen könnte, anstatt sie zu stärken.

Tamay Besiroglu versucht, diese Bedenken auszuräumen, indem er erklärt, dass das wirtschaftliche Wohlergehen nicht nur auf Arbeiterlöhne angewiesen ist, und schlägt vor, dass die Menschen sich stattdessen auf Mieten, Dividenden und staatlich bereitgestellte Sozialprogramme verlassen könnten – vorausgesetzt natürlich, dass KI-generierte Gewinne schließlich in Steuern umschlagen, die solche Programme finanzieren!

Während **Tamays Besiroglus** Vision umstritten und bei weitem nicht realisiert wird, sind die Herausforderungen, die er angehen will, real und allgegenwärtig.

Die heutigen KI-Agenten erweisen sich immer noch häufig als unzuverlässig, kämpfen mit komplexen Aufgaben und es fehlt an der Fähigkeit, kohärente Operationen über längere Zeiträume aufrechtzuerhalten.

Mechanize erkennt diese erheblichen Einschränkungen an, während er sich selbst als einer der vielen Player vorschlägt – neben Technologieriesen wie **Salesforce**, **Microsoft** und **OpenAI** – und arbeitet aktiv daran, die Autonomie und Zuverlässigkeit in der digitalen Assistententechnologie zu verbessern.



Tamay Besiroglu

Unabhängig von den philosophischen Debatten, die durch ihre kühnen Ziele aufgeworfen werden, tritt *Mechanize* in eine bereits überfüllte Arena von Unternehmen ein, die bestrebt sind, die Technologie von KI-Agenten zu nutzen, die sich über verschiedene Spezialisierungen – von der Vertriebsautomatisierung bis hin zur Finanzanalyse – erstreckt.

Tamay Besiroglus umstrittene Verfolgung bleibt ein weiteres Signal dafür, dass das Silicon Valley weiterhin mutige und potenziell disruptive Visionen der kommenden Welt vorantreiben wird. Vorerst, inmitten von Kontroversen und Kontrolle, treibt *Mechanize* weiterhin aggressiv voran und rekrutiert aktiv. 🐼

Die Ziele von *Mechanize*

Wie plant Mechanize, die aktuellen Einschränkungen von KI-Modellen anzugehen?
Die Betreiber von Mechanize haben sich das ehrgeizige Ziel gesetzt, die **vollständige Automatisierung der Wirtschaft durch KI** zu erreichen.

Dafür müssen sie die heutigen Grenzen bestehender KI-Modelle – wie zum Beispiel mangelnde Robustheit, geringes Kontextverständnis, beschränktes Planungsvermögen und fehlende Fähigkeit zu autonomem Handeln über längere Zeiträume – gezielt überwinden.

Mechanize kombiniert innovative Simulationsumgebungen, praxisnahe Benchmarks und fortgeschrittene Lernmethoden, um die aktuellen Schwächen heutiger KI zu überwinden – mit dem Fernziel, die vollständige Automatisierung von Wissensarbeit und Büroarbeit zu ermöglichen. Kritiker sehen darin sowohl eine technische als auch eine gesellschaftliche Herausforderung, doch der Ansatz gilt als einer der ambitioniertesten in der KI-Industrie.

Die Strategie von Mechanize lässt sich wie folgt zusammenfassen:

1. Simulation komplexer Arbeits- umgebungen

Mechanize plant, virtuelle Arbeitsumgebungen (Simulationsplattformen) zu schaffen, in denen KI-Agenten praktisch alle Schritte und Abläufe typischer Büro- oder Wissensarbeit realitätsgetreu üben und optimieren können. Dies geht über reine Sprachmodelle hinaus und ermöglicht den Agenten Feedbackschleifen,

Fehlerkorrekturen und das Lernen aus komplexen Interaktionen – ähnlich wie Menschen am Arbeitsplatz.

2. Verstärkendes Lernen (Reinforcement Learning) und Benchmarks

Statt KI nur mit Beispielen zu füttern, setzt Mechanize auf **Reinforcement Learning**: Ihre Agenten erhalten Belohnungen für erfolgreiche Aufgaben und lernen aus Fehlern.

Um Fortschritt objektiv messbar zu machen, entwickelt das Unternehmen Benchmarks, die reale Arbeitsszenarien abbilden (im Gegensatz zu klassischen KI-Benchmarks wie Multiple-Choice oder kurze Aufgaben).

3. Iterative Leistungssteigerung durch praxisnahe Tests

Durch die Kopplung von Simulation, Benchmarking und RL-Training kann Mechanize schrittweise immer schwierigere Aufgaben automatisieren. Die Modelle lernen dadurch, nicht nur einzelne Aufgaben zu erfüllen, sondern über längere Zeitabläufe hinweg zu planen, zu improvisieren und mit unvorhergesehenen Ereignissen umzugehen.

4. Integration von multimodalen Fähigkeiten

Langfristig wird Mechanize nicht nur auf Textverständnis setzen, sondern auch Bild-, Code- und weitere Eingabetypen einbinden. Dadurch könnten KI-Agenten beispielsweise Dokumente analysieren, Präsentationen erstellen oder sogar einfache Programmieraufgaben lösen.

5. Mensch-KI-Kollaboration als Zwischenziel

Kurz- bis mittelfristig sieht das Konzept von Mechanize sogenannte “Human-in-the-Loop”-Prozesse vor: Menschen prüfen, korrigieren und lenken die KI, wodurch die Systeme kontinuierlich verbessert und zuverlässiger werden. 🐼

John McCarthy, Miterfinder der modernen KI

John McCarthy (1927 – 2011) war ein US-amerikanischer Logiker, Informatiker und Autor. Erfinder der Programmiersprache **LISP**. Für seine bedeutenden Beiträge zur **Künstlichen Intelligenz** erhielt McCarthy 1971 den **Turing Award** und 1988 den Kyoto-Preis.

John McCarthy wurde 1927 in Boston als Sohn irischer und litauischer Einwanderer geboren. Während der Depression zog die Familie viele Male um und kam schließlich in Los Angeles an, wo sein Vater Organisator einer Arbeitergewerkschaft wurde und seine Mutter in der Frauenwahlrechtsbewegung aktiv war.

Obwohl er mit einer Krankheit behindert war, war John McCarthy ein außergewöhnlich helles Kind. Er erhielt 1948 den Bachelor of Science im Fach Mathematik vom *California Institute of Technology*. Den Doktorgrad erwarb er drei Jahre später – 1951 – an der *Princeton University* bei Solomon Lefschetz (*Projection operators and partial differential equations*).

1951 promovierte er in Mathematik in Princeton. Über Dartmouth und am MIT wurde er 1962 ordentlicher Professor in *Stanford*, wo er bis zu seiner Pensionierung blieb.

Nach Aufgehalten an der *Stanford University*, am *Dartmouth College* und im *Massachusetts Institute of Technology* (MIT) war er ab 1962 Professor in Stanford, wo er 2000 emeritierte.

Er war bis zu seinem Tod als Professor der Stanford University aktiv und kommentierte das Weltgeschehen oft in Internet-Foren aus einer mathematisch-wissenschaftlichen Perspektive.



John McCarthy wurde als ein Mensch der "auf die Zukunft fokussiert war" beschrieben, er "erfetzte, erfand, und erfand". In den 1960er Jahren konzipierte er die Idee des *Time-Sharing*, welches es Computer-Benutzern ermöglichte, Daten zu teilen, indem sie mit einem zentralen Computer verknüpft waren; dies senkte schließlich die Kosten für die Verwendung von Computern. Diese Innovation wurde als wesentlicher Beitrag zur Entwicklung des *Internets* und als Vorläufer des *Cloud-Computings* beschrieben – eine Methode, um Daten auf einem über das Internet zugänglichen Remote-Server zu speichern.

Die Dartmouth-Konferenz

Während seines Aufenthalts in Dartmouth organisierte John McCarthy die bahnbrechende Konferenz in Sachen *Künstliche Intelligenz*. Dort traf er *Marvin Minsky* (*), der einer der führenden Theoriker auf dem Feld der KI wurde.

(*) Biografie von Marvin Minsky: KI-Newsletter Nr. 8



John McCarthy gewann ein Stipendium für die Universität MIT und Minsky, der in Harvard war, schloss sich ihm an. 1959 gründeten beide das *MIT Artificial Intelligence Laboratory*. Ihre Ansichten fingen jedoch an zu divergieren und John McCarthy kehrte nach Stanford zurück. Er gründete das *Stanfords KI-Labor*, bekannt als *SAIL*, das zum Rivalen des MIT wurde. Von 1965 bis 1980 war er dessen Direktor.

In den 1960er und 1970er Jahren spielte das *Stanford-Labor* eine entscheidende Rolle bei der Schaffung der Systeme, die viele menschliche Fähigkeiten nachahmen – Vision, Zuhören, Denken und Bewegung.

Manchmal präsentierte er Erfindungen und lud den Homebrew Computer Club, eine Hobbygruppe aus dem Silicon Valley, ein, sich in den Stanford-Labors zu treffen. Die Gruppe umfasste zwei von Apples Gründermitgliedern, Steve Jobs und Steven Wozniak.

In den 1970er Jahren präsentierte McCarthy ein Papier über den Kauf und Verkauf von Computern, in dem er den E-Commerce voraussah. John McCarthys erster Doktorand war *Raj Reddy*, der später ebenfalls den Turing Award gewonnen hat. McCarthy war auch Doktorvater von Turingpreisträgerin *Barbara Liskov*.

In den Jahren von 1957 bis 1959 wurde er als Sloan Fellow im Bereich Physical Science gefördert.

John McCarthy ist der Erfinder der Programmiersprache LISP, dessen Design er 1960 in der Fachzeitschrift *Communications of the ACM* vorstellte. LISP war eine der ersten Implementierungen eines Logikkalküls auf einem Computer.

Auszeichnungen

Er gehörte der American Academy of Arts and Sciences (1974), der American Association for the Advancement of Science, der American Mathematical Society und der Association for Computing Machinery an.

1987 wurde er in die National Academy of Engineering und 1989 in die National Academy of Sciences aufgenommen.

Von 1983 bis 1984 war er Präsident der *American Association for Artificial Intelligence* und er gehörte seit 1975 dem Editorial Board der Zeitschrift *Artificial Intelligence Journal* an. 🐼

1956: Die erste Konferenz über KI

Die erste Konferenz über Künstliche Intelligenz, die *Dartmouth Conference*, fand im Sommer 1956 statt. In dem Förderantrag an die Rockefeller Foundation für diese Konferenz prägte John McCarthy 1955 den Begriff "Künstliche Intelligenz."

Die Dartmouth Conference gilt als Geburtsstunde der Künstlichen Intelligenz als Forschungsgebiet. Beantragt, geplant und durchgeführt wurde sie von John McCarthy, Marvin Minsky, Nathaniel Rochester (IBM) und Claude Shannon.

Unter dem vollständigen Namen *Dartmouth Summer Research Project on Artificial Intelligence* fand sie im Sommer 1956 vom 19. Juni bis zum 16. August am Dartmouth College in Hanover, New Hampshire statt.

Allen Newell, wichtiger KI-Pionier

Allen Newell (1927–1992) spielte eine zentrale Rolle bei der Entwicklung der Künstlichen Intelligenz; er ist einer ihrer Pioniere. Seine Beiträge waren grundlegend für die Fortentwicklung dieses Feldes...

Wie bringt man einen Computer dazu, die Denkprozesse des Menschen nachzuahmen? Diese Frage beschäftigt Allen Newell seit den 1950er Jahren und wurde zu seiner Lebensaufgabe.

Seine Forschungen bilden die Grundlage für die *Künstliche Intelligenz*, die uns heute überall im Alltag begegnet. Newells Ansatz, Künstliche Intelligenz-Forschung mit dem Verständnis menschlicher Denkprozesse zu verbinden, prägt nach wie vor die KI-Forschung. Seine Arbeiten finden sich noch immer in Psychologie-Lehrbüchern zum Thema *Problemlösen*.

Als Allen Newell 1927 in San Francisco geboren wird, sind Computer noch eine Art Science-Fiction. Der Sohn eines Radiologen entspricht nicht dem Bild eines Nerds: Er ist naturverbunden; und als Jugendlicher will er Ranger in der Sierra Nevada werden...

Doch dann erwacht seine Liebe zur Wissenschaft – durch die umstrittenen Kernwaffentests der USA im Südpazifik. Newells Vater ist einer der Forscher, die im Bikini-Atoll Kernwaffen testen und beschäftigt seinen Sohn Allen als Assistenten. Allen wollte unbedingt “einer von ihnen sein und an interessanten Problemen arbeiten.”

Dann studiert Allen Newell Mathematik und Physik. In seinem ersten Job beschäftigt er sich mit der Frage, wie es zu Entscheidungsfindungen kommt. Bei einem Vortrag über die noch junge Computerwissen-



schaft findet Newell sein “interessantes Problem”, das ihn sein Leben begleiten wird: “Es war der Gedanke, dass wir mit mechanistischen Mitteln verstehen konnten, wie der menschliche Verstand funktioniert.”

Newell will das menschliche Gehirn und die Denkprozesse künstlich nachahmen, um sie zu erforschen. Die so entstehende Künstliche Intelligenz ist sozusagen das Nebenprodukt der philosophischen Entschlüsselung des Geistes.

Langjährige Forschung mit Nobelpreisträger Herbert A. Simon

Allen Newell entwickelte zusammen mit Herbert A. Simon einige der frühesten und einflussreichsten KI-Programme:

1. Logic Theorist

Logic Theorist (1956): Dies war das erste Programm, das in der Lage war, mathematische Theoreme zu beweisen. Konkret führte der *Logical Theorist* den Beweis von 38 Theoremen aus den

Principia Mathematica von Bertrand Russell und Alfred North Whitehead.

Dieses Ergebnis war ein Meilenstein der Künstlichen Intelligenz, da gezeigt wurde, dass Programme zu Aktionen fähig sind, für die ein Mensch Intelligenz braucht.

2. General Problem Solver

General Problem Solver (GPS) (1957): Dieses Programm konnte ein breiteres Spektrum von Problemen lösen, darunter Theorembeweise und Spiele wie Schach oder die Türme von Hanoi.

3. Kognitive Architektur

Newell entwickelte die kognitive Architektur *Soar*, die als Versuch einer einheitlichen Theorie menschlicher Kognition gilt. *Soar* integriert Erkenntnisse der Kognitionspsychologie in ein Computerprogramm, um menschliches Verhalten zu modellieren und vorherzusagen.

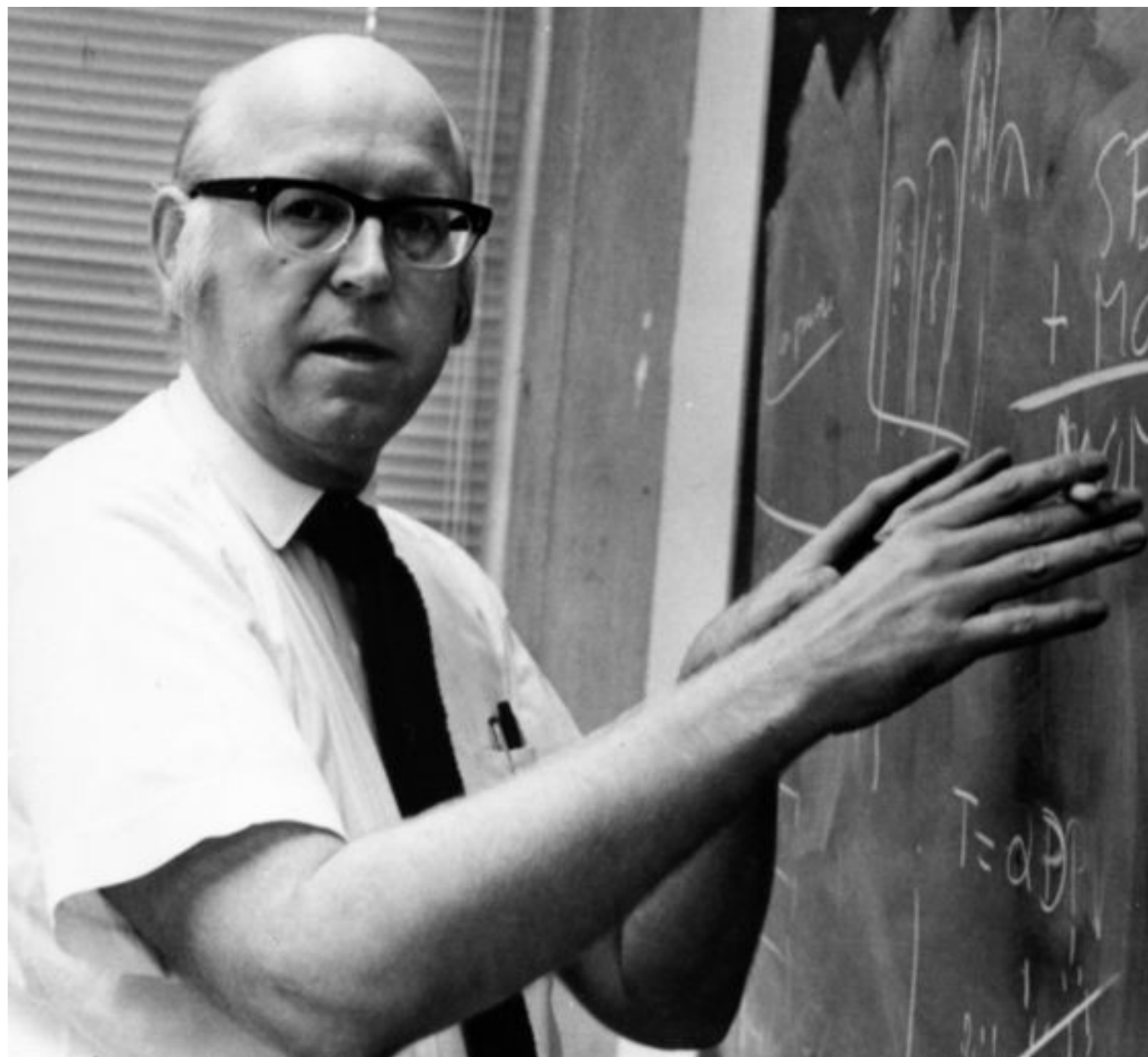
Auszeichnungen und Anerkennung

Für ihre Forschungen zur Programmierung und Künstlichen Intelligenz wurden Allen Newell und Herbert Simon 1975 mit dem *Turing-Preis* ausgezeichnet, den höchsten Preis der Computerwissenschaften. Im Jahr darauf erhält Newell eine Professur an der *Carnegie Mellon University* in Pittsburgh.

Hier forscht Newell weiter an dem Problem, das ihn seine gesamte Laufbahn beschäftigt: der *Natur des menschlichen Geistes*. Als sein letztes Projekt wählt Allen Newell bewusst eine Mammutaufgabe, die er nicht abschließen können würde – eine Theorie, die alle kognitiven Leistungen beschreibt und simulierbar machen soll.

Wunschgemäß überdauert ihn sein letztes Projekt. Allen Newell stirbt 1992 mit 65 Jahren an Krebs. Eine einheitliche Kognitionstheorie ist bis heute nicht gefunden. Dabei findet Newells Verbindung von Künstlicher Intelligenz und menschlicher Denkprozesse auch in den aktuellen Forschung zur Künstlichen Intelligenz noch immer Beachtung.

Weitere Auszeichnungen: Wahl in die American Academy of Arts and Sciences (1972) und die National Academy of Sciences. Allen Newells Beiträge haben die Grundlagen für viele moderne KI-Technologien gelegt und seinen Platz als einer der Väter der Künstlichen Intelligenz gefestigt. 🐼



Newells Forschung konzentrierte sich auf die Frage, wie der menschliche Verstand funktioniert und wie man dieses Verständnis nutzen kann, um intelligente Systeme zu entwickeln. Er verband dabei Erkenntnisse aus der Informatik mit denen aus der Psychologie.

Herbert Simon, Verhaltensforscher

Herbert Alexander Simon (1916 – 2001) war ein US-amerikanischer Sozialwissenschaftler. Im Jahr 1978 erhielt er den Nobelpreis für den Bereich Wirtschaftswissenschaften. Er erforschte das menschliche Verhalten in Organisationen und war KI-Pionier.

Seine interdisziplinäre Forschung erstreckte sich über die Bereiche Kognitionswissenschaft, Informatik, öffentliche Verwaltung, Management und Politikwissenschaft. Neben seiner akademischen Arbeit war Herbert A. Simon auch als Berater tätig; er diente zwischen 1968 und 1971 als wissenschaftlicher Berater für die Präsidenten Johnson und Nixon.

Herbert A. Simon hinterließ ein beeindruckendes Erbe als Pionier in mehreren wissenschaftlichen Bereichen und als einflussreicher Denker, der die Art und Weise, wie wir menschliche Entscheidungsfindung und künstliche Intelligenz verstehen, grundlegend verändert hat.

Herbert Simon war Sohn deutschstämmiger Eltern. Sein Vater **Arthur Simon** war in die USA ausgewandert; seine Mutter Edna Merkel, eine Pianistin, hatte Vorfahren in Prag und Köln.

Nach der Schulzeit an einer High School strebte er dem Vorbild eines Onkels, einem Sozialwissenschaftler, nach und studierte ab 1933 in Chicago Sozialwissenschaften mit mathematischem Schwerpunkt; er strebte an, die Gesellschaftswissenschaft durch eine analytische und systematische Basis zu stützen.

Im Jahr 1936 schloss er sein Studium ab und war bis 1939 als Forschungsassistent in der Stadtverwaltung und von 1939 bis 1942 als Leiter einer Forschungsgruppe an der *Universität Berkeley* tätig.



Herbert Simon

Sein Thema war bereits seit seiner Universitätszeit die **Entscheidungsfindungsprozesse in Organisationen**. Parallel zu seiner Zeit in Berkeley promovierte er an der *University of Chicago* über Entscheidungsfindung in Verwaltungsorganisationen.

Herbert Simon trat 1942 in Chicago am *Illinois Institute of Technology* eine Stelle als Politologe an, wobei er parallel Kontakte zu Ökonomen (u. a. Milton Friedman) an der *University*

of Chicago aufbaute und Vorlesungen in Wirtschaftswissenschaften hörte.

Zu seiner mathematisch-sozialwissenschaftlichen Ausrichtung kam nun eine fundierte Ausbildung in Wirtschaftswissenschaften. Seit 1949 baute er gemeinsam mit anderen Wissenschaftlern am *Carnegie Institute of Technology* einen Aufbaustudiengang für Industrieverwaltung auf. Und er wurde Mitglied der Kommission, die für die Koordination des Marshall-Plans verantwortlich war.

Das ursprüngliche Forschungsgebiet der **Entscheidungsfindung** innerhalb von Organisationen hatte Simon nie ganz verworfen. Das Thema beschäftigte ihn kontinuierlich über die nächsten Jahrzehnte.

Der 'Logical Theorist'

Seit 1954 erforschte Herbert Simon diese Prozesse mit Hilfe von Computersimulationen, und er entwickelte 1956 mit **Allen Newell** (Seite 28) den *Logical Theorist*. Dieses Programm war erstmals dazu in der Lage, eine Menge von logischen Theoremen zu beweisen.

Ein weiterer wichtiger Schritt hin zum maschinellen Problemlösen war die ebenfalls mit Newell entwickelte Software '*General Problem Solver*' (Allgemeiner Problemlöser).

1959 wurde Herbert Simon in die American Academy of Arts and Sciences sowie die American Philosophical Society gewählt, 1967 in die National Academy of Sciences. Daneben war er zwischen 1968 und 1971 im wissenschaftlichen Beraterstab der Präsidenten Johnson und Nixon tätig.

Im Jahr 2001 starb Herbert Simon im Alter von 84 Jahren in Pittsburgh, Pennsylvania.

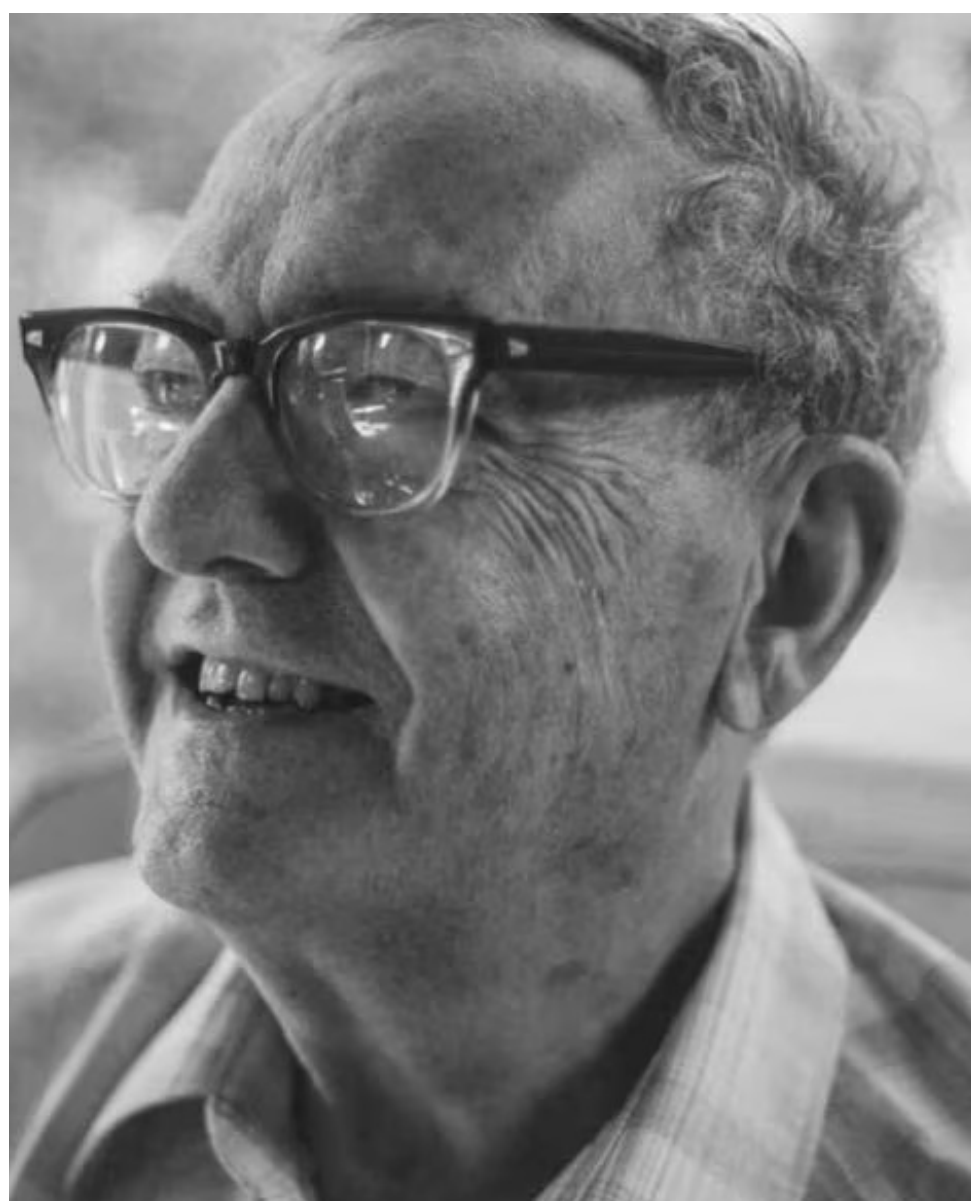
Auszeichnungen

1975 Turing Award (für die Arbeiten mit Newell)

1978 Alfred-Nobel-Preis für Wirtschaftswissenschaften

1988 John-von-Neumann-Theorie-Preis

1995 IJCAI Award for Research Excellence



Herbert Simon

Herbert Simon leistete bahnbrechende Beiträge in verschiedenen Disziplinen:

Er entwickelte die Konzepte der "begrenzten Rationalität" und des "Satisficing", wofür er 1978 den Nobelpreis für Wirtschaftswissenschaften erhielt.

Zusammen mit Allen Newell entwickelte er 1956 den *Logical Theorist* und später den *General Problem Solver*, zwei wegweisende KI-Programme. Für seine Arbeit in diesem Bereich wurde er 1975 mit dem Turing Award ausgezeichnet.

Simon untersuchte menschliche Problemlösungsprozesse und trug wesentlich zur Entwicklung der *kognitiven Psychologie* bei.

Er erforschte Entscheidungsprozesse in Organisationen und revolutionierte das Verständnis von administrativem Verhalten.

Veröffentlichungen

Simon war ein äußerst produktiver Wissenschaftler, der 27 Bücher und fast tausend Artikel verfasste. Seine wichtigsten Veröffentlichungen:

Administrative Behavior. A Study of Decision-Making Processes in Administrative Organizations. New York/London, 1947.

Das Verwaltungshandeln. Eine Untersuchung der Entscheidungsvorgänge in Behörden und privaten Unternehmen. Stuttgart 1955; Übersetzung der 3. Auflage: *Entscheidungsverhalten in Organisationen. Eine Untersuchung von Entscheidungsprozessen in Management und Verwaltung.* Verlag Moderne Industrie, Landsberg 1981.

mit James G. March: *Organizations.* 1958; Blackwell, Malden (Mass.) 1993.

Organisation und Individuum. Menschliches Verhalten in Organisationen. Betriebswirtschaftlicher Verlag Gabler, Wiesbaden 1976.

The Shape of Automation: for Men and Management. Evanston, New York 1965.

Perspektiven der Automation für Entscheider. Schnelle, Quickborn 1966.

The Sciences of the Artificial. The MIT Press, Cambridge (MA) 1969.

Die Wissenschaften vom Künstlichen. Kammerer & Unverzagt, Berlin 1990.

Reason in Human Affairs. Stanford University Press, 1983

Homo rationalis. Die Vernunft im menschlichen Leben. Campus-Verlag, Frankfurt/New York 1993.

Models of my life. Basic Books, New York 1991.

Literatur über Simon

David Klahr, Kenneth Kotovsky: A Life of the Mind: Remembering Herb Simon. In: APS Observer. April 2001

Crowther-Heyck Hunter: Herbert A. Simon. The bounds of Reason in Modern America. Johns Hopkins Univ. Press, Baltimore, 2005. 🐼



Intelligence

explosion?