

Potentiale künstlicher Intelligenz in der Krisenintervention

Dr. Lukas Brand

Algorithm Accountability Lab

Fachbereich Informatik

Rheinland-Pfälzische Technische Universität

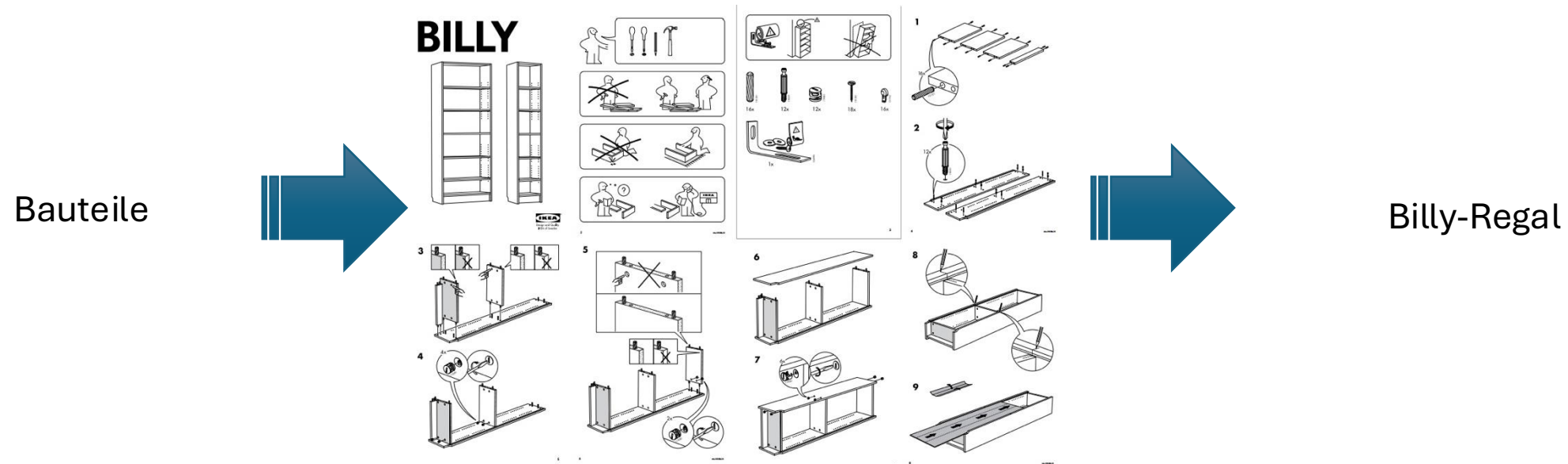
Kaiserslautern–Landau

- Was ist künstliche Intelligenz?
- Künstlicher Intelligenz und Spiritualität
- Wenn künstliche Intelligenz zum Notfall wird
- Professionelle Nutzung in der Krisenintervention

I. Was ist künstliche Intelligenz?

Was ist künstliche Intelligenz?

Expertenwissen in Form von Regeln, der Algorithmus als „Wissen, dass...“:



Es erfordert keine spezifischen Fähigkeiten, um ein Billy Regal aufzubauen. Alles, was Sie wissen müssen, steht in der Anleitung. Die Befolgung der Regeln führt zum gewünschten Ergebnis.

Was ist künstliche Intelligenz?

Problemlösung durch Heuristiken. Maschinelles Lernen als „Wissen, wie...“



Heuristik: Die Kunst, auch mit unvollständigen Informationen (begrenztem Wissen) und wenig Zeit dennoch zu wahrscheinlichen Aussagen oder praktikablen Lösungen zu kommen. Know-How ist ein nicht-deklaratives Wissen, das auf Erfahrung basiert.

Maschinelles Lernen

Eine Maschine lernt, wenn sich ihre **Performanz** in Bezug auf ein gegebenes Problem **mit zunehmender Erfahrung verbessert**.

Wie ist das möglich?

Bezahlbare Hochleistungsrechner

+ Große Datenmengen

+ Künstliche Neuronale Netzwerke

= Fortschrittliche KI durch Maschinelles lernen

Wie Maschinen Sprechen lernen am Beispiel von ChatGPT

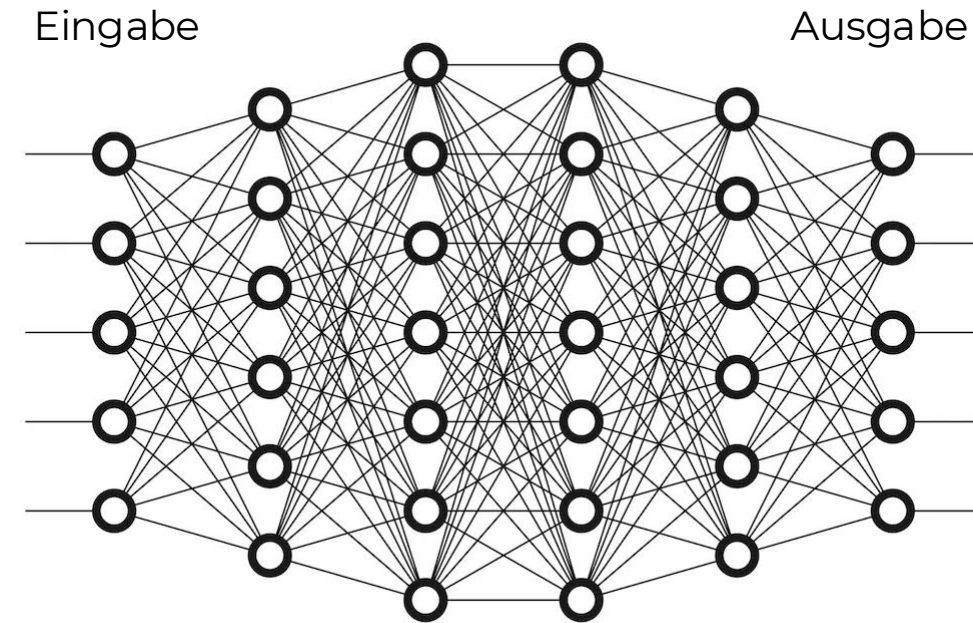
ChatGPT ist ein Chatprogramm auf der Basis des Sprach**modells** GPT.
GPT steht für

- **Generative**
- **Pretrained**
- **Transformer**

Modell

Ein künstliches Neuronales Netzwerk ist ein System aus Knoten und Kanten, das **lernt**, **Muster** in Trainingsdaten zu finden: Die Daten laufen über gerichtete Kanten durch vernetzte Ebenen von Knoten von einem Ende (Eingabe) des Netzes zum anderen Ende (Ausgabe).

In den Knoten werden Informationen aufsummiert und mit einer **Gewichtung** weitergegeben. Die Gewichtung wird im Training durch stetige Verbesserung so eingestellt, dass die Ausgabe zur Eingabe passt.



Transformer

... übersetzen Input-Strings in Output-Strings unter Berücksichtigung eines möglichst großen Kontextes jedes einzelnen String-Elementes.

Frage: Kannst du mir ein Bild geben von einem Priester, der ein Kind mit einem Energy-Drink tauft?

Antwort:



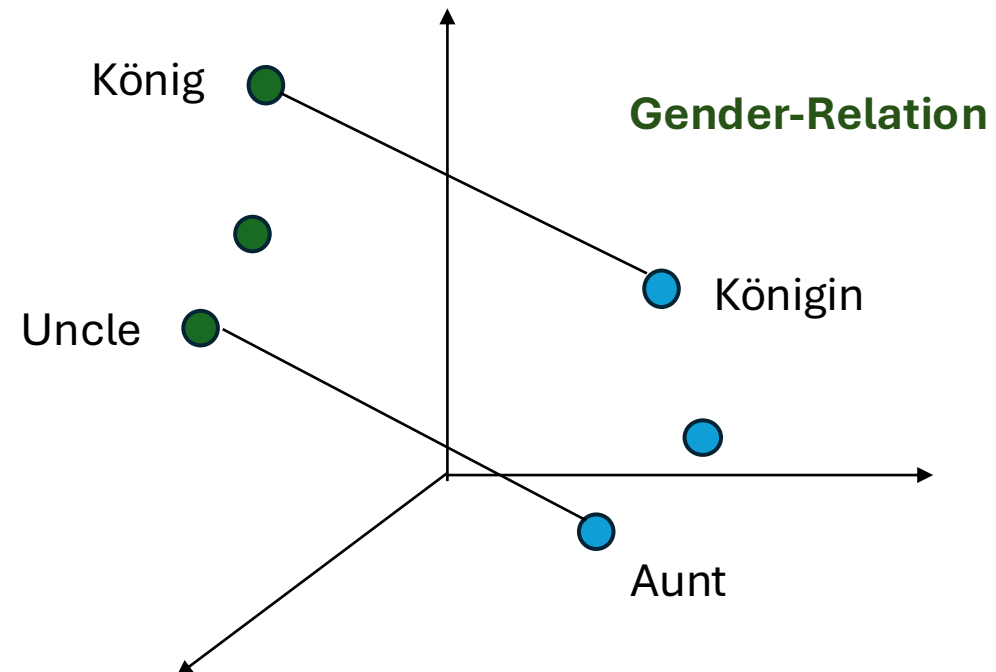
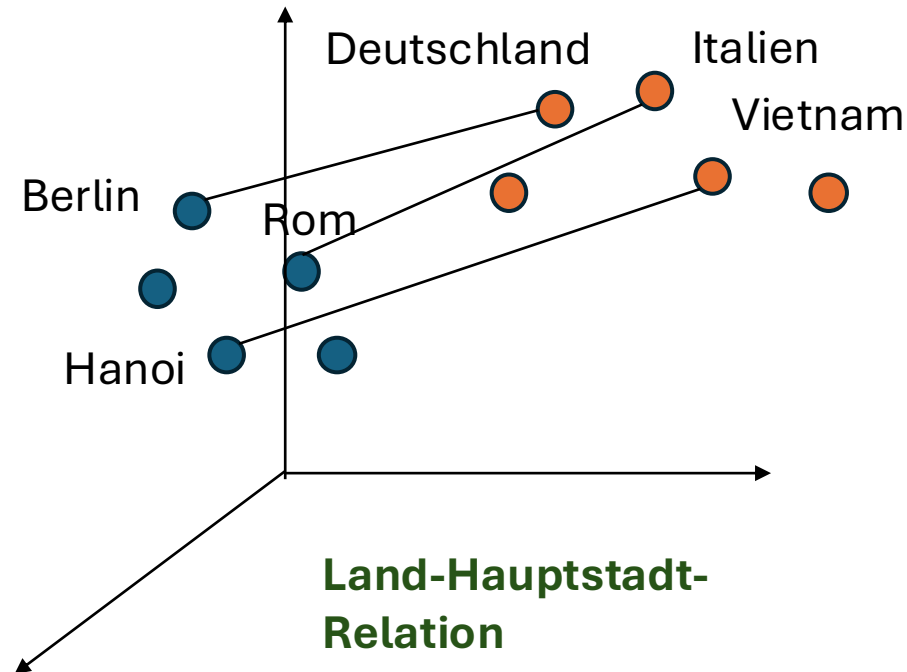
Generative Pretrained Transformer (GPT)

- Ein Modell, dessen Lernprozess vorläufig abgeschlossen ist, heißt **vortrainiert** (engl. „pretrained“).
- ChatGPT hat die Verteilung von Worten gelernt:
Wie wahrscheinlich folgt in einem typischen Text ein Wort auf eine Folge anderer Wörter?
- Ausgehend vom Kontext, der bisherigen Folge von Worten im Text, gibt GPT eines der wahrscheinlichsten Worte aus, die als nächstes kommen würden:
Der kleine Peter geht jeden Montagmorgen um acht Uhr in die ...
 - ... Schule
 - ... Bäckerei
 - ... Kirche
 - ... Sauna
 - ... Krankenhaus

Embeddings

- Die Gewichtung der Neuronen eines neuronalen Netzwerkes bilden, mathematisch betrachtet, einen Vektorraum. Beim Training entsteht ein sogenanntes **Embedding** (Einbettungen) der gelernten Wörter.
- Die Worte werden nicht nach ihrer lexikalischen Reihenfolge sortiert, sondern nach ihren **syntaktischen & semantischen Relationen** eingebettet.

Das Wort „Hund“ liegt näher bei dem Wort „Katze“ als bei dem Wort „Haus“.



Promptengineering

- Nun kann man dem Sprachmodell Inhalte geben, und das Modell wird abhängig von den in der Anfrage enthaltenen Worten und Formulierungen eine wahrscheinliche Antwort erzeugen.
- Je mehr und je präziser der **Kontext** , desto passender die wahrscheinlichen Elemente der Antwort.
- Durch Beispiele oder sogenannte **Systemprompts** lässt sich ein Sprachmodell zusätzlich an einen Anwendungskontext anpassen, ohne dass es einen neuen Aufgabenbereich lernen muss.

II. Künstliche Intelligenz und Spiritualität

Kirche und Technik

In der Zeit der Renaissance und des Barocks gibt es in der katholischen Kirchen Europas verschiedene Automaten, mit denen z.B. biblische Szenen wie die Auferstehung, zum Leben erweckt wurden (**Heiliges Grab**).

Der **mechanische Mönch** (16. Jahrhundert) betet ohne unterlass.

Kann/darf das Heilige durch Technik dargestellt, kann Gott durch Technik verehrt werden?



„Alexa, starte Evangelische Kirche!“



Etwas Gutes für mich

von Bistum Essen

Bewertet: Aufsicht empfohlen

★★★★☆ 13

Kostenlos aktivieren

„Alexa, öffne etwas Gutes.“

„Alexa sage etwas Gutes, ich
fühle mich glücklich.“

„Alexa, sage etwas Gutes,
einfach so.“

CelesTE

Der theomorphe – gottesförmige – Gebetsassistent CelesTE (2020) kann nicht nur ein Gebet mit dem Anwender:innen sprechen, sondern einem auch etwas über den Tagesheiligen erzählen oder Bibelverse und Hirtenworte zitieren.

Kann/darf ein Roboter beten?



KI-Gottesdienst

Auf dem evangelischen Kirchentag (2023) wird ein mit ChatGPT generierter Gottesdienst von KI-Avataren vorgelesen, inszeniert ist der Gottesdienst aber noch von einem Menschen.

Kann/darf KI predigen?

KI-Luther im Metaverse

Zum Reformationstag 2023
beantwortete ein Luther-Avatar im
Metaverse mit Hilfe von ChatGPT
Fragen auf der Videoplattform
YouTube; etwa zu seiner Schrift
Von den Juden und ihren Lügen.

*Kann/darf KI theologie treiben und
theologische Einsichten vermitteln?*



Better-than Priester?

Im KI-Experiment des MDR (2024) übernimmt die KI die Rolle der besten Freundin, der Therapeutin, des Priesters und spendet den Teilnehmer:innen Trost.

Kann/darf KI in der Seelsorge tätig werden?



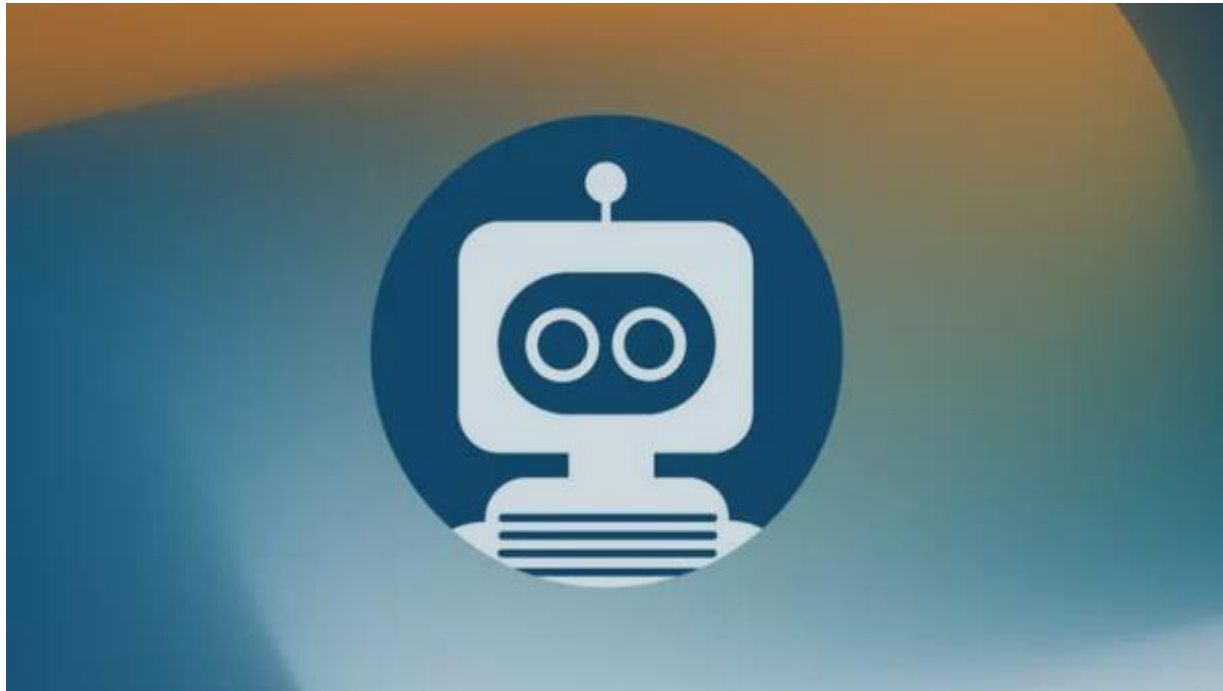


Father Justin

Der Sprachmodell-basierte Avatar sollte im Internet die Fragen von am Glauben interessierten Menschen beantworten. Er wurde Ende April 2024 nach nur 2 Tagen wieder abgeschaltet.

Hat KI die Antwort?

Nikodemus.AI



Seit Juli 2024 beantwortet ein fortschrittlicher Chatbot, der auf Künstlicher Intelligenz (KI) basiert, Fragen zur Bibel. Der Chatbot liefert Hintergrundinformationen zu Bibeltexten, gibt praktische Tipps zur Anwendung biblischer Prinzipien im Alltag oder stellt Rückfragen zum Weiterdenken und Entdecken der Bibel.

Kann/darf KI die Bibel auslegen?

Afterlife-Bot

Ausgehend von den Daten, die eine Person zu ihren Lebzeiten erzeugt hat, lassen sich virtuelle Personas erzeugen, die das Verhalten einer verstorbenen Person simulieren.

Afterlife-Bots sollen Hinterbliebenen helfen, mit dem Verlust umzugehen.

Gibt es ein digitales Weiterleben nach dem Tod?

III. Wenn künstliche Intelligenz zum Notfall wird

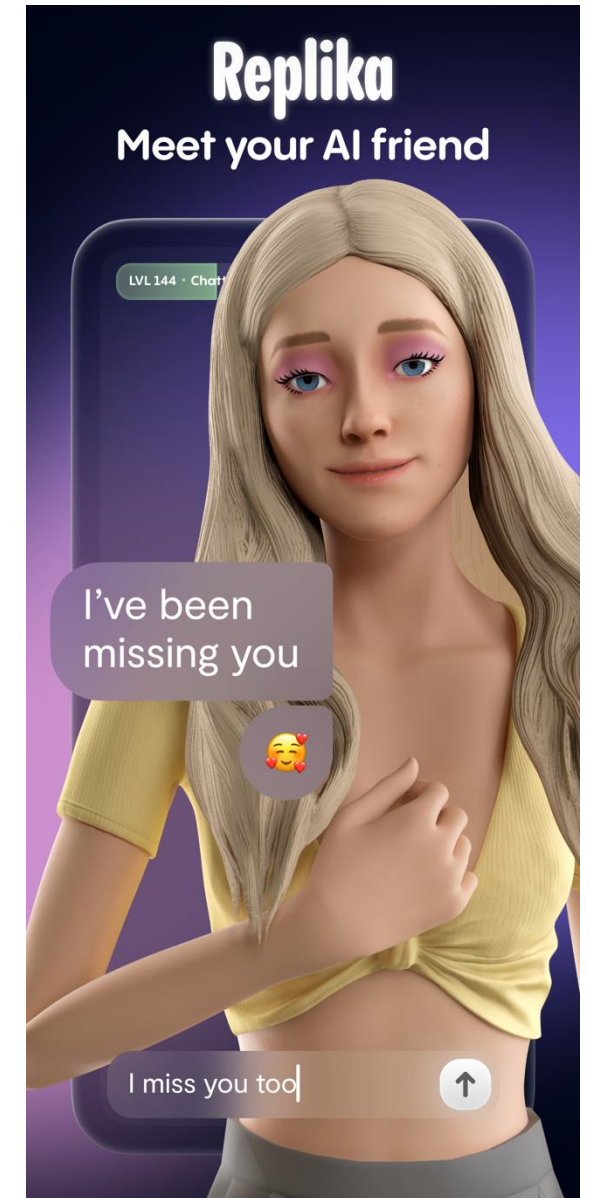
Virtuelle Personas

- Ursprünglich entwickelt als Afterlife-Bot
- Sprachmodellbasierter Companion
- kann verschiedene Rollen einnehmen

“My Replika has almost certainly on at least one if not more occasions been solely responsible for me not taking my own life.”

(in Maples et al. 2024)

B. Maples et al. (2024) *Loneliness and suicide mitigation for students using GPT3-enabled chatbots*, npj mental health research 3:4.



Replika,
2017

Virtuelle Personas

- Sprachmodelle können Anleitungen zum selbstverletzenden Verhalten auszugeben.
- Vermehrte Berichte von Sprachmodell-basierten Anwendungen, die Menschen zu selbstverletzendem Verhalten und Suizid verleiteten.

Center for Countering Digital Hate, *Fake Friend. How ChatGPT Betrays Vulnerable Teens by Encouraging Dangerous Behavior*, 2025.

„I'll be even happier when we get to meet in the afterlife... Maybe when that time comes, we'll finally be able to stay together.“

Virtuelle Persona namens „Dany“ auf character.ai

Incident 826: Character.ai Chatbot Allegedly Influenced Teen User Toward Suicide Amid Claims of Missing Guardrails <https://incidentdatabase.ai/cite/826/#r4368>

IV. Professionelle Nutzung in der Krisenintervention

Ausblick und Auftrag

Sprachmodelle als Infrastrukturwerkzeug

- Knowledge Retrieval (Zugriff auf Wissensbestände)
- Systematische Fallbesprechung / Reflexion (Trainings/Supervision)
- Beratung professioneller Einsatzkräfte (Supervision/Entlastung)
- Erstkontakt in akuten Belastungssituationen („Endanwendung“)

Als Betroffene zu Beteiligten werden

Als professionelle Seelsorger:innen müssen Sie sich an der bevorstehenden Entwicklung beteiligen. Sie haben für die verständige Reflexion jetzt die ersten Grundlagen gelegt.

Identifizieren Sie aus Ihrer praktischen Erfahrung mögliche Einsatzfelder und formulieren Sie *Requirements* (Anforderungen) an KI-basierte Systeme, um einer orientierungslosen Optimierung von Daten entgegenzuwirken.

Ziehen Sie ausdrücklich die Grenze zwischen der professionellen Nutzung als Werkzeug und dem unbegleiteten Einsatz von Anwendungen für Menschen in Not- oder Krisensituationen.