Conversational AI. Dialogsysteme, Chatbots, Assistenten

Veranstalter: Christoph Ringlstetter

Sitzung 0: Organisation, Vorstellung, Einordnung, industrielle Anwendung

Was machen wir denn heute.

- Orga. Referate, Zulassung, Termine.
- Vorstellung der Veranstaltung
- Diskussion zur Ausrichtung
- Besprechung Paper zum Paradigmshift
- Glimpse Anwendungen von GenAI in der Industry und AI Teams in der industriellen Organisation

Voraussetzungen Schein:

Referat 20 -30 Minuten zu einem der Themen oder zu einem Programmierprojekt.
Literatur wird bereitgestellt kann aber auch selbst vorgeschlagen werden.
Vorbesprechung wird angeboten – vermutlich Samstag

Ausarbeitung des Themas ca. 15 Seiten oder Abgabe des Programmierprojektes

Programmierprojekt Beispiel Rasa Paper. TOD.

Mögliche Themen und Schwerpunkte des Seminars:

- ➢Introduction to Large Language Models (LLMs) -User's perspective
- ➢Language models and beyond
- ➢Architecture engineering and scaling law -Transformer and beyond
- ➢Training LLMs from scratch -Pre-training, SFT, learning LLMs with human feedback
- ➢Efficiency in LLMs. Environmental perspective.
- ➢Prompt engineering
- ➢Knowledge engineering and reasoning
- ➢Multimodal LLMs
- ➢LLMs in vertical domains
- ➢Tools and large language models
- ➢Privacy, bias, fairness, toxicity and holistic evaluation
- ➢Alignment and limitations

Mögliche Themen und Schwerpunkte des Seminars:

- ➢Hyperscaler Frameworks: Google, Microsoft, Amazon
- ➢Langchain Framework
- ➢Huggingface and selfhosting of an LLM.
- ➢Societal implications
- ➢Artificial Emotions
- ➢Memory emulation
- ➢RAGs
- ➢LLMs and GenAI in Industry
- ➢Philosophical perspective on AGI
- ➢Weekly News on GenAI
- ➢Action Models Autonomous Agents
- ➢Models worldwide

Möglicher Seminarplan – Erster Wurf

Semester 14. Oktober – 07. Februar

Phase I:

17.10. **Vorstellung der Teilnehmer:innen. Überblick zum Seminar, Kursverlauf. Industrielle Anwendungen.**
24.10. **Einführung zu Conversatioal AI. History. Status.**
31.10. **LLMs as Basis for new stage of Conversational AI.**
07.11. **Practical overview: Langchain, LLM access.**

Möglicher Seminarplan – Erster Wurf

Semester 14. Oktober – 07. Februar

Phase II:

14.11. **Prompting, Emerging Abilities, Models Overview**

21.11. **Transformers in Depth, Training LLMs., RLHG**

28.11**. RAGs, Grounding/Halunzination, Evaluation of LLMs**

05.12. **Multimodal Models, Vertical LLMs and Bots**

12.12. **Ethical considerations, Bias**

19.12. **Societal impact including sustainability, strategical access**

Möglicher Seminarplan – Erster Wurf

Semester 14. Oktober – 07. Februar

Phase III:

09.01. **Langchain advanced, Hyperscaler Frameworks, Hugging Face**

16.01. **Reasoning, Knowledge engineering**

23.01. **Memory, Perspective on AGI**

30.01. **Emotions, Speech**

06.02. **Wrap up - Reserve**

Literaturhinweise.

Bücher:

Jurafsky&Martin

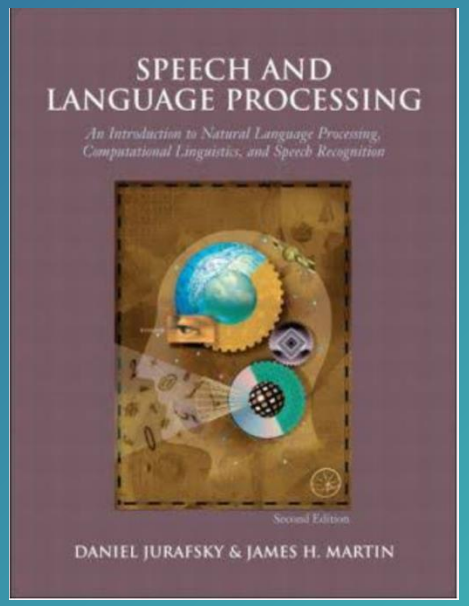
Langchain Book,

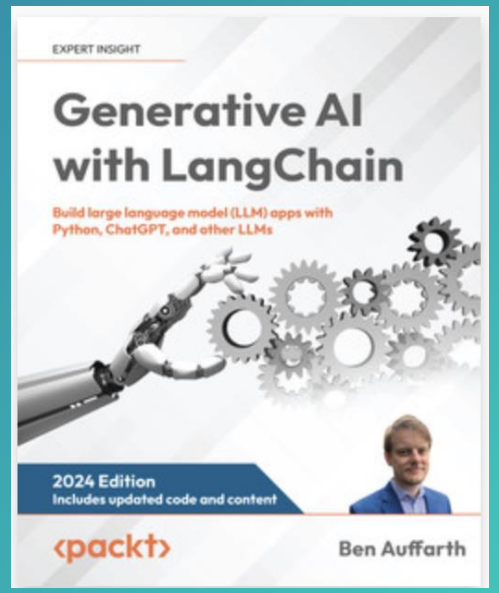
Transformers Book

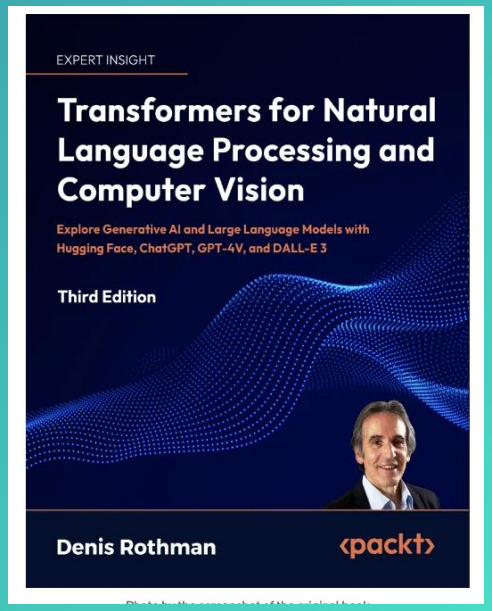
Survey Papers

Spezielle Papers

Andere Quellen:







- Kurse: Stanford, U HK, Shenzhen.
- Workshops:
- Konferenzen: ACL, Coling, NIPS, Workshop on Spoken Dialogue Systems
- News Feeds: New York Times AI Pod Cast, Daily Zap
- Blogs der Decisive Players: z.B. OpenAI Blog, Deep Mind Blog, Baidu?

OpenAI Status o1

Open AI Blog.

https://openai.com/o1/

Kurz Film anschauen

September 20 · Research

Top row (left to right): Mark Chen, Giambattista Parascandolo, Trapit Bansal, Łukasz Kaiser, Hunter Lightman, Karl Cobbe, Łukasz Kondraciuk, Szymon Sidor, Noam Brown, Hongyu Ren, Liam Fedus, Hyung Won Chung

Bottom row (left to right): Ilge Akkaya, Jakub Pachocki, Shengjia Zhao, Jason Wei, Wojciech Zaremba, Jerry Tworek

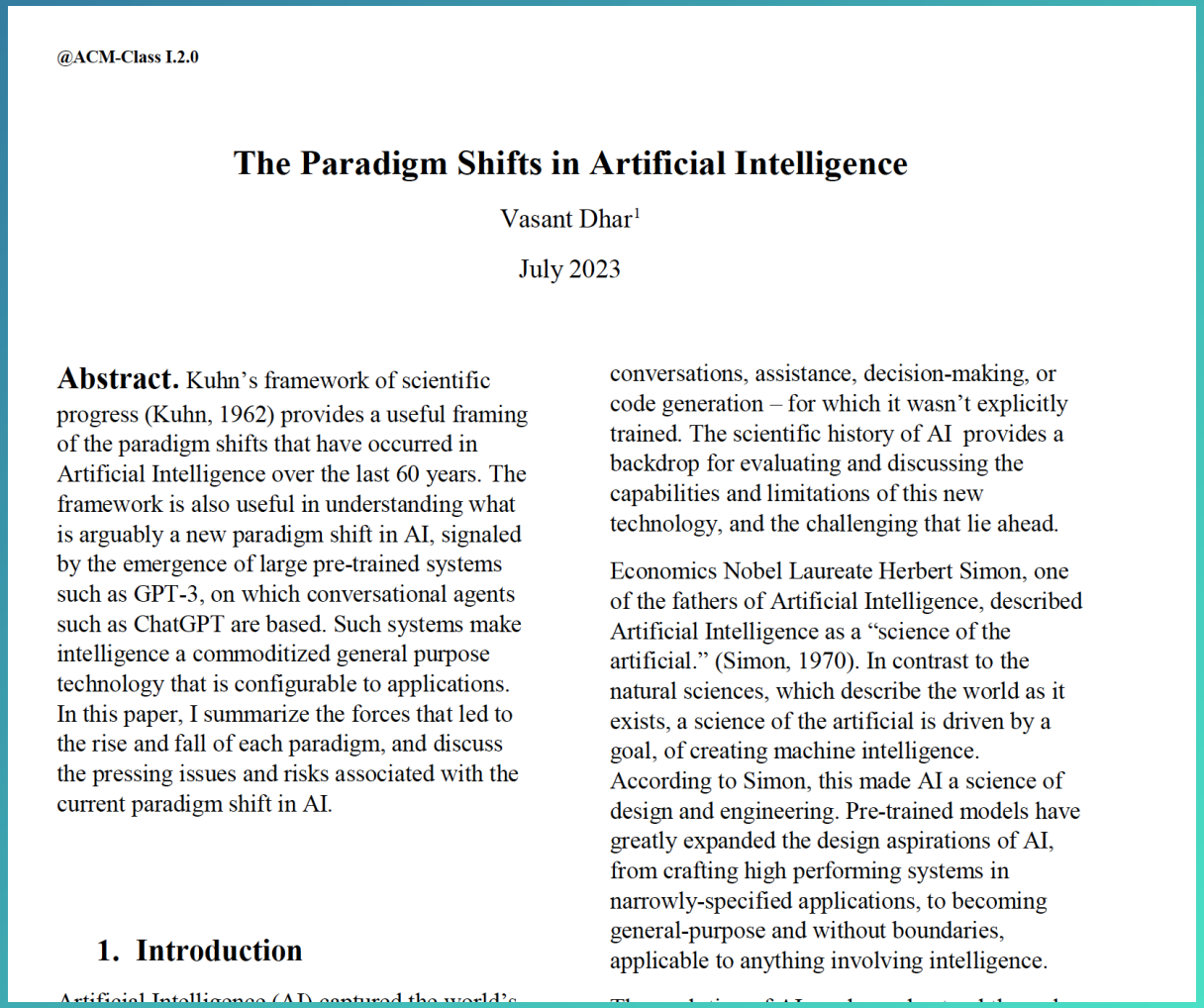
Besprechung Paper Paradigmshift.

- „The evolution of AI can be understood through Kuhn's (1962) theory of scientific progress in terms of "**paradigm shifts**." A paradigm is essentially a set of theories and methods accepted by the community to guide inquiry. It's a way of thinking. Kuhn describes science as a process involving occasional "r**evolutions**" stemming from crises faced by the dominant theories, followed by periods of "normal science" where the details of the new paradigm are fleshed out."
- From:
  The Paradigm Shifts in Artificial Intelligence Vasant Dhar1 07/23
-

„AI is transitioning from a bespoke set of tools to a "**General Purpose Technology,**" from which applications are assembled. Like electricity, intelligence becomes a commodity."

configurable, pervasive, improving, innovational

But: Sollow citation on IT revolution. "IT was everywhere except in the productivity statistics." (Solow, 1987).



@ACM-Class I.2.0

The Paradigm Shifts in Artificial Intelligence

Vasant Dhar[1]

July 2023

Abstract. Kuhn's framework of scientific progress (Kuhn, 1962) provides a useful framing of the paradigm shifts that have occurred in Artificial Intelligence over the last 60 years. The framework is also useful in understanding what is arguably a new paradigm shift in AI, signaled by the emergence of large pre-trained systems such as GPT-3, on which conversational agents such as ChatGPT are based. Such systems make intelligence a commoditized general purpose technology that is configurable to applications. In this paper, I summarize the forces that led to the rise and fall of each paradigm, and discuss the pressing issues and risks associated with the current paradigm shift in AI.

1. Introduction

Artificial Intelligence (AI) captured the world's

conversations, assistance, decision-making, or code generation – for which it wasn't explicitly trained. The scientific history of AI provides a backdrop for evaluating and discussing the capabilities and limitations of this new technology, and the challenging that lie ahead.

Economics Nobel Laureate Herbert Simon, one of the fathers of Artificial Intelligence, described Artificial Intelligence as a "science of the artificial." (Simon, 1970). In contrast to the natural sciences, which describe the world as it exists, a science of the artificial is driven by a goal, of creating machine intelligence. According to Simon, this made AI a science of design and engineering. Pre-trained models have greatly expanded the design aspirations of AI, from crafting high performing systems in narrowly-specified applications, to becoming general-purpose and without boundaries, applicable to anything involving intelligence.

Auftakt.

- ChatGPT: wir konnten zum ersten mal mit einer künstlichen Installation über alle möglichen Themen wie mit einem Menschen sprechen.
- In einer Vielzahl von Aufgaben für die es nicht explizit trainiert wurde kompetent: Konversation, NLP, Decission making, Science. Prompting.
- Simon: AI als Wissenschaft des künstlichen -- Design und Engineering
- GPT haben das Design von eng spezifizierten Anwendungen hin zu generellen Anwendungen von Intelligenz verschoben.
- Erklärungsversuch mit Kuhns Theorie vom Paradigmenwechsel: Paradigma als Menge von Theorien, Methoden: Way of Thinking
- In AI shift von symbolischer Repräsentation von Erklärungen zu direkter Modellierung durch die Daten

Paradigmenwechsel in AI: Expertensyteme 1980s

- Expertensysteme attraktiv in engen gut beschriebenen Domänen mit identifizierbarem und definierbarem menschlichem Wissen: Gesundheit, Ingeneurswesenen
- Unsicherheit als Problem: statistische Modellierung.
- Insgesamt: Engineering Bottleneck, Regelexplosion, statisch rigid.
- „Human reasoning and language seemed much too complex and heteroge nous to be captured by top-down specification of relationships. Progress stalled, as the reality, both in research and practice, fell short of expectations.“

Paradigmenwechsel: Machine Learning. 1990s

- Automatisches lernen von Regeln aus Daten, geleitet durch menschliche Intuition: Feature Engineering: Bayes, SVMs,
- Modelle gelernt durch Verlustfunktion (MLE): breiter Einsatz auch in Data Science. Vorhersage von Krankheiten, Wetter, Q Ereignissen
- Funktionsapproximation. Maschine als Hypothesentester
- Entspricht auch dem Wissenschaftstheoretischen Zielbild: falsifizierbare Theorien. Popper
- Auch schon NNs aber als intransparent, hoher Rechenaufwand
- Dennoch: neues Bottleneck, data curration, labeling, maschinenverständlich
- Ziel: „ingest raw data such as numbers, images, notes or sounds directly, ideally without curation by humans“

Deep Learning 2010s

- “Deep learning,” made abig dent in the feature engineering bottleneck by providing a solution to perception, such as seeing, reading, and hearing“
- (Hinton, 1992; LeCun and Bengio, 1998). Stacked Layers die die Unzulänglichkeiten von Vorgängerarchitekturen überwanden: allgemeiner Funktionsapproximator mit der Eigenschaft: abstrakte Konzepte nahe am Outputlayer werden duch Basis Features aus den Layern näher am Inputlayer erzeugt (composed)
- „In language modeling, for example, the core learning task is typically to predict the next occurrence of an input sequence. This requires a considerable amount of knowledge and understanding of the relationships among the different parts of the input. Large language models use a special configuration of the “Transformer” neural architecture, which represents language as a contextualized sequence, where context is represented by estimating dependencies between each pair of the input sequence (Vaswani et.al 2017).

Deep Learning

- Im Gegensatz zu Expertensystemen wo das Wissen in lokalisierbaren Chunks liegt ist es beim DL über die Gewichte verteilt und viel schwerer interpretierbar - Explainability
- „Bowman (2023) conjectures that the autocomplete task was serendipitous: it was just at the right level of difficulty, where doing well conversationally forced the machine to learn a large number of other things about the world.“

	Knowledge Source	Exemplar	Capability	Data Curation
Expert Systems	Human	Rules	Follows	High
Machine Learning	+ Databases	Rules/Networks	+ Discovers Relationships	Medium
Deep Learning	+ Sensory	Deep Neural Networks	+ Senses Relationships	Low
General Intelligence	+ Everything	Pre-Trained Deep Neural Networks	+ Understands the World	Minimal

Table 1: The Paradigm Shifts in AI

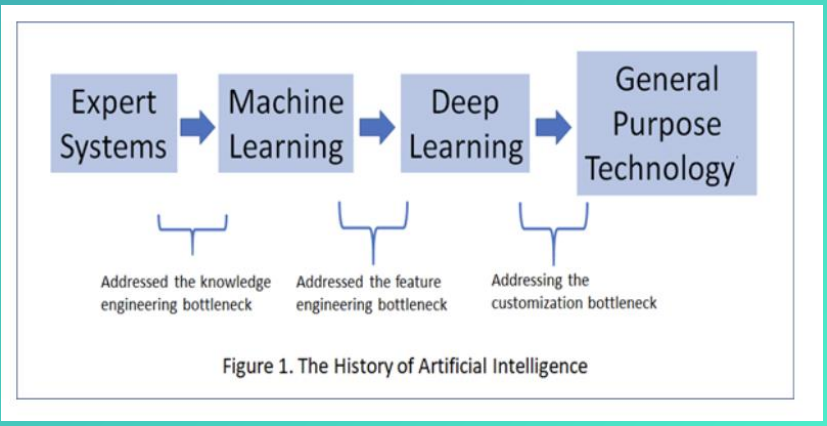


Figure 1. The History of Artificial Intelligence

AGI

- Jedes Paradigma erweiterte die Anwendungssphären
- Pretrained Models sind domänen- und applikationsunabhängig: Internet als Datenquelle
- Komunikation in natürlicher Sprache mit der Maschine als entscheidendem Durchbruch
- Prinzipien: mit zunehmender Gößе neue Fähigkeiten, Vorhersage des nächsten Wortes scheint einen harten Zusammenhang mit Intelligenz zu haben.
- „ The first paradigm was “Learn from humans." The next one was “Learn from curated data." This was followed by “Learn from any kind of data." The current paradigm is “Learn from all kinds of data it in a way that transfers to novel situations.”

AI als General Purpose Theory

- LLMs und conversational AI haben eine neue Welle von Anwendungen auf industriellem Level ermöglicht. Disruptiv.
- Wie Elektrizität wird Intelligenz zur Commodity
- Eigenschaften: “pervasive – they are used as inputs by many downstream sectors), inherent potential for technical improvements, And innovational complementarities – the productivity of R&D in downstream sectors multiplies as a consequence of innovation in the general purpose technology, creating productivity gains throughout the economy.
- AI ist in den frühen Stadien: nicht makroproduktiv nur early adopter

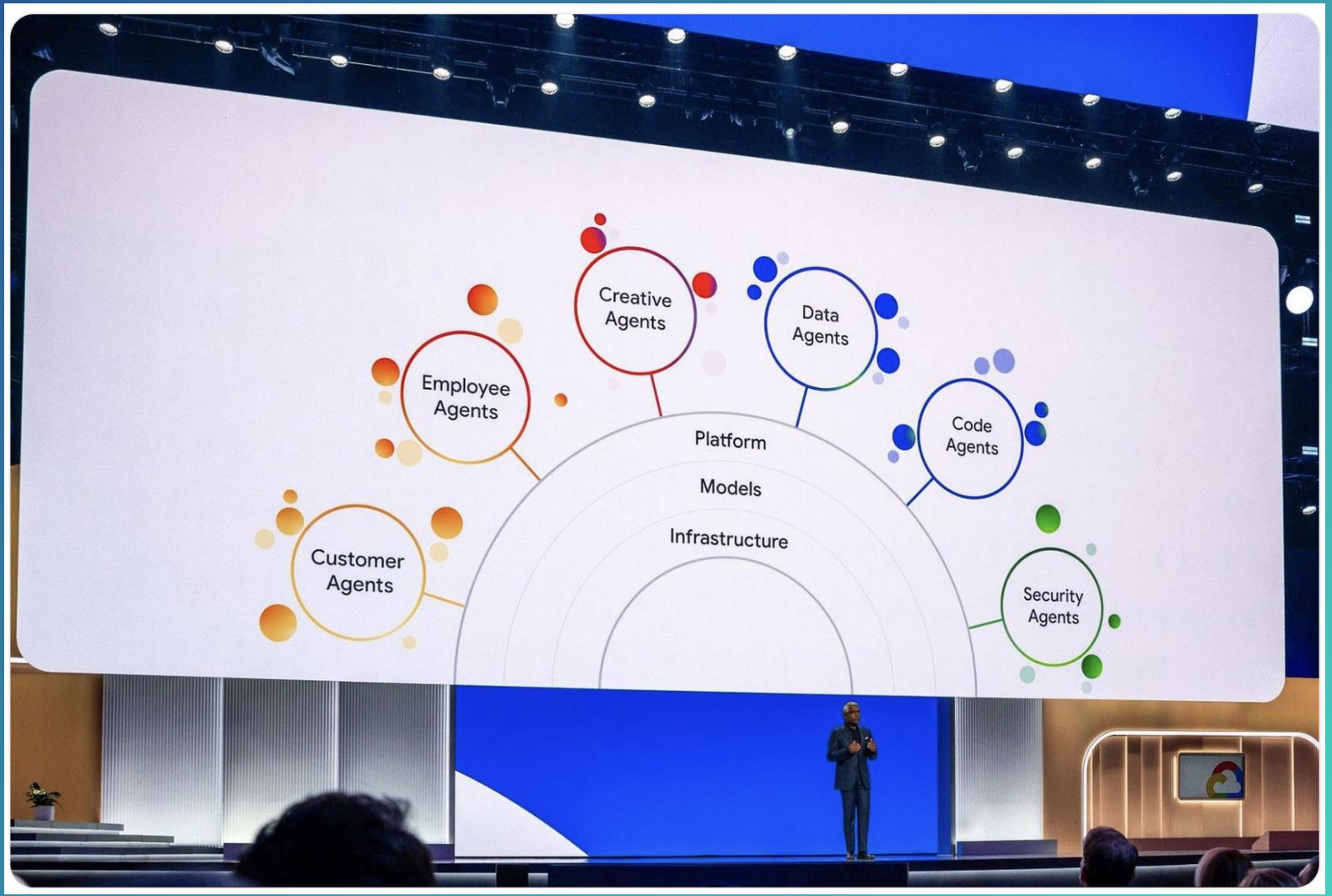
Probleme und Herausforderungen

- Natürlich dynamisch: noch am Anfang des Paradigmenwechsels
- Opaque dadurch auch Widerstand: EU Gesetzgebung
- Korrektheit nicht beweisbar
- Beherrschbarkeit der Maschine nicht gegeben. Indeterministisch. Unvorhersehbar.

AI, such as “save the planet,” we have no idea about the sub-goals the machine will create in order to achieve its larger goals. This is known as “the alignment problem,” in that it is impossible to determine whether the machine’s hidden goals are aligned with ours. In saving the planet, for example, the AI might determine that humans pose the greatest risk to its survival, and hence they should be contained or eliminated

Industrieller Praxisglimpse

Google Beispiel Usecases, Volkswagen Data:lab. NLP, DL, CV



https://cloud.google.com/transform/101-real-world-generative-ai-use-cases-from-industry-leaders

Durch AI Agenten Verbesserung der Produktivität, automatisieren von Prozessen, neue Customer Experience: Schlüsselbereiche: customer service; employee empowerment; code creation; data analysis; cybersecurity; and creative ideation and production.

Zuhören, Bedürfnisse verstehen, Produkte/Services empfehlen. Ohne Brüche über Kanäle hinweg: Internet, Mobile, physischer/virtueller Salespunkt. Integration mit Voice und Video.

- ING Bank aims to offer a superior customer experience and has developed a gen-AI chatbot for workers to enhance self-service capabilities and improve answer quality on customer queries. Watch the session to learn more.
- Volkswagen of America built a virtual assistant in the myVW app explore their owners' manuals. See how it works
- Mercedes Benz will infuse e-commerce capabilities into its online storefront with a gen AI-powered smart sales assistant.
- Snap has deployed the multimodal capability of Gemini within its "My AI" chatbot and has since seen over 2.5-times as much engagement within Snapping to My AI in the United States. Watch the video.

https://cloud.google.com/transform/101-real-world-generative-ai-use-cases-from-industry-leaders

Employee agents helfen den Kolleginnen produktiver zu sein und besser zusammenzuarbeiten. Prozesse straffen, Routineaufgaben erledingen, Anfragen der Beschäftigten erledigen. Bei der Kommunikation helfen. Zusammenfassen, Formulieren, Editieren, Übersetzen.

Dun & Bradstreet built an email-generation tool with Gemini that helps sellers create tailored, personalized communications to prospects and customers for its research services.

England’s Football Association is training Vertex AI on the FA's historical and current scouting reports so they can be transformed into concise summaries, helping national teams discover future talent.

Hemominas, Brazil's largest blood bank, partnered with Xertica to develop an omnichannel chatbot for donor search and scheduling, streamlining processes and enhancing efficiency. The AI solution has the potential to save half-a-million lives annually by attracting more donors and optimizing blood supply management.

Randstad, a large HR services and talent provider, is using Gemini for Workspace across its organization to transform its work culture, leading to a more culturally diverse and inclusive workplace that’s seen a double-digit reduction in sick days. Watch the video.

Bank of New York Mellon built a virtual assistant to help employees find relevant information and answers to their questions. Watch the session to learn more.

Code agents design, code, test und operate. Double-digit gains in productivity.

Labelbox has built a fully managed AI model evaluation solution directly integrated into the Vertex AI platform, allowing Google Cloud users to seamlessly launch human evaluation jobs and set specific criteria for evaluation, such as question-answering and summarization; this eases and accelerates the ability to deploy human-in-the-loop AI systems with higher levels of trust and authority.

Magic is building a developer platform with a 100-million-token context window, so organizations can upload extremely large code bases and more easily query and build on them using gen AI assistance.

Wayfair piloted Code Assist, and those developers with the code agent were able to set up their environments 55 percent faster than before, there was a 48 percent increase in code performance during unit testing, and 60 percent of developers reported that they were able to focus on more satisfying work. Watch the [illegible] to learn more.

https://cloud.google.com/transform/101-real-world-generative-ai-use-cases-from-industry-leaders

Data agents ersetzen z.T. Data analysts and Researcher. Frage zu internen und externen Quellen, neue Modelle entwickeln. Ideenbrowsing.

Addy AI is helping mortgage lenders and banks automate their lending processes with custom AI models trained on Vertex AI. For example, the platform can extract loan opportunity details from lengthy email threads with numerous attachments.

Bayer Crop Science has developed Climate FieldView, a comprehensive agricultural platform with more than 250 layers of data and billions of data points; AI-powered recommendations allow farmers to design and monitor their fields for greater yields and efficient fertilization, with the added benefit of reduced carbon emissions. The Asteroid Institute is using AI to discover hidden asteroids in existing astronomical data. This is a major focus for astronomers researching the evolution of the Solar System, investors and businesses hoping to fly missions to asteroids, and for all of us who want to prevent future large asteroid impacts on Earth. Watch the session to learn more.

134. Contextual is working with Google Cloud to offer enterprises fully customizable, trustworthy, privacy-aware AI grounded in internal knowledge bases.

Security agents erhöhen extrem die Geschwindigkeit von Untersuchungen, automatisieren Monitoring, Verbessern Respons. Können auch Cyberattacs wie Prompt Injection bekämpfen.

Apex Fintech is using Gemini in Security to accelerate the writing of complex threat detections from hours to a matter of seconds.

Palo Alto Networks is using Gemini to create a grounded AI assistant for 24/7 security platform support in order to improve agent efficiency and response time; grounding the assistant in organizational data and security protocols has greatly improved the accuracy of responses.

163. BBVA uses AI in Google SecOps to detect, investigate, and respond to security threats with more accuracy, speed, and scale. The platform now surfaces critical security data in seconds, when it previously took minutes or even hours, and delivers highly automated responses.

https://cloud.google.com/transform/101-real-world-generative-ai-use-cases-from-industry-leaders

https://cloud.google.com/transform/101-real-world-generative-ai-use-cases-from-industry-leaders

Creative Agents in Design and Produktion Bilder Slides. Konzepte. Marketing, audio and video Produktionsteams „With creative agents, anyone can become a designer, artist, or producer.“

Adore Me marketers write differentiated product descriptions in one hour, a tedious task which used to take 30-40 hours a month thanks to Gemini for Google Workspace.

Puma is using Imagen to customize product photos on its website, saving time and ensuring they are locally relevant across markets; PUMA India has already seen a 10% increase in click through rate. Watch a demo.

Radisson Hotel Group personalized its advertising at scale in collaboration with Accenture and using Vertex AI and Gemini models, training them on extensive datasets stored in BigQuery; ad teams saw productivity rise around 50% while revenue increased from AI-powered campaigns by more than 20%

Paramount currently relies on manual processes to create the essential metadata and video summaries used across its Paramount+ platform for showcasing content and creating personalized experiences for viewers. VertexAI Text Bison is now helping to streamline this process. Watch the session to learn more.

https://cloud.google.com/transform/101-real-world-generative-ai-use-cases-from-industry-leaders