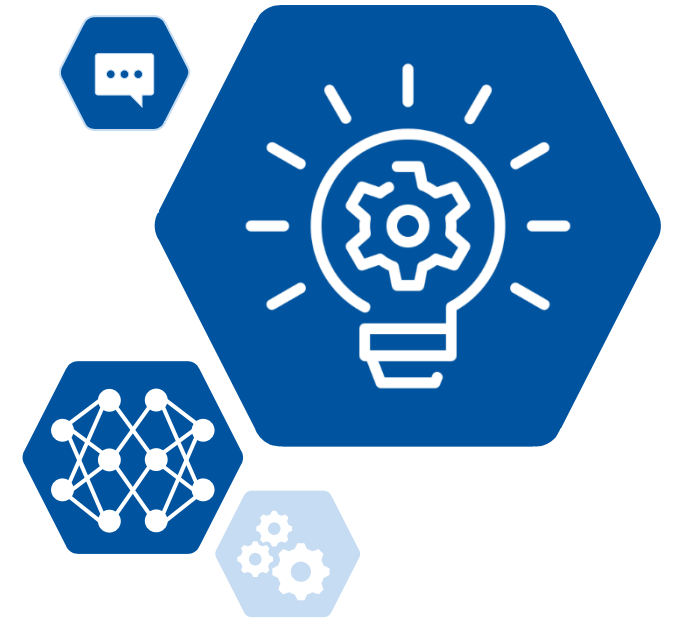
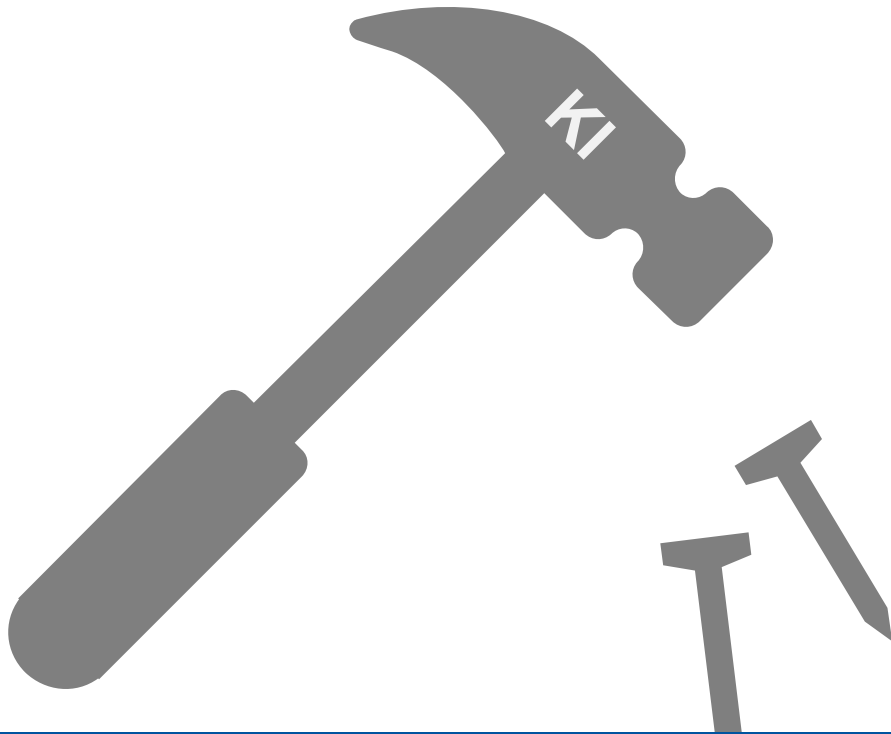




Wenn das tolle neue Werkzeug ein Hammer ist, sieht jedes Problem aus wie ein Nagel !

LLMs in Lehre, Forschung und Software-Entwicklung

GI-Themenabend 10.03 Köln

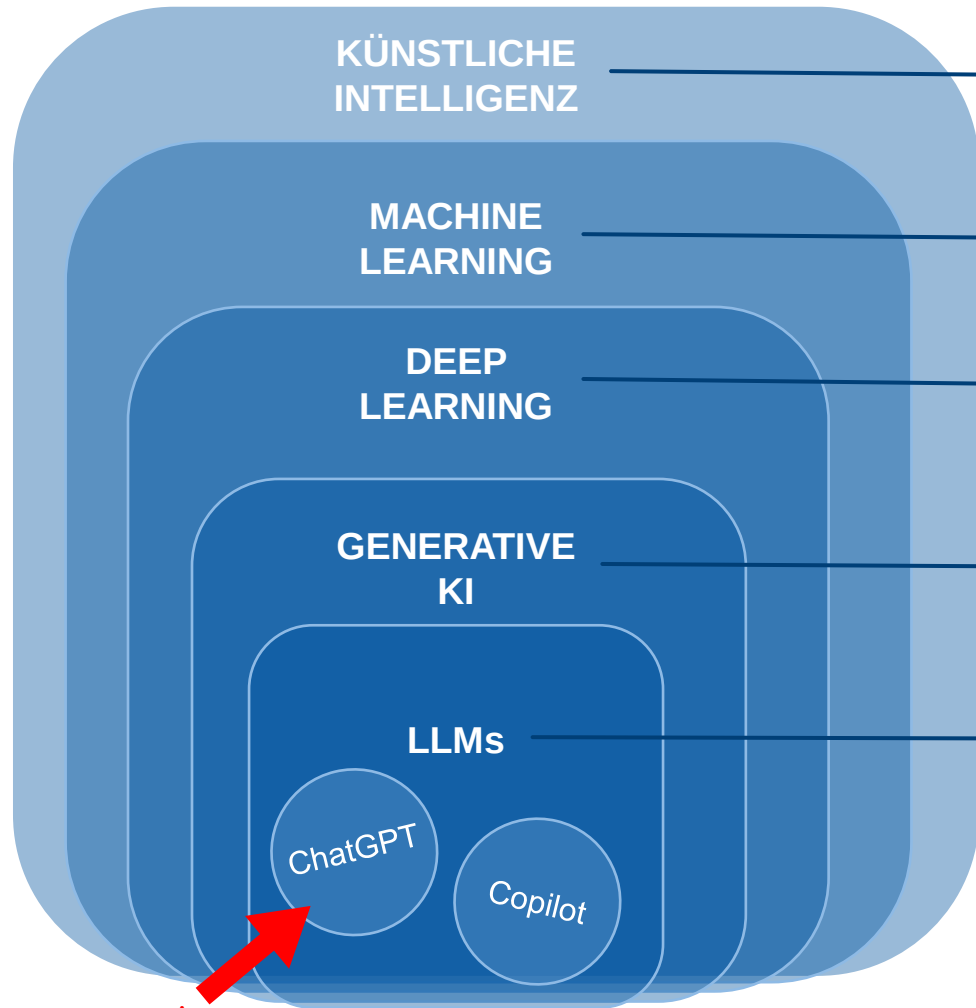
Dr. Lukas Netz
Software Engineering
RWTH Aachen
se-rwth.de



SE
Software
Engineering

RWTH AACHEN
UNIVERSITY

Sammelbegriff KI



KÜNSTLICHE
INTELLIGENZ

Forschungsgebiet, dass sich mit der Entwicklung **intelligenter Maschinen** befasst, die wie Menschen arbeiten und reagieren.

MACHINE
LEARNING

Statistische Algorithmen, die aus Daten lernen und auf ungesehene Daten verallgemeinern können.

DEEP
LEARNING

Teilgebiet von Machine Learning, dass **Neuronale Netze** nutzt, um komplexe Muster in Daten zu erkennen.

GENERATIVE
KI

Eine Art von KI, die in der Lage ist neue **Inhalte zu erzeugen**, indem sie bestehende Muster und Daten analysiert (genAI).

LLMs

Große **Spachmodelle** sind fortschrittliche KI-Modelle, die **große Mengen an Textdaten** verarbeiten, um menschenähnliche Texte zu generieren oder zu verstehen.

ChatGPT

Copilot

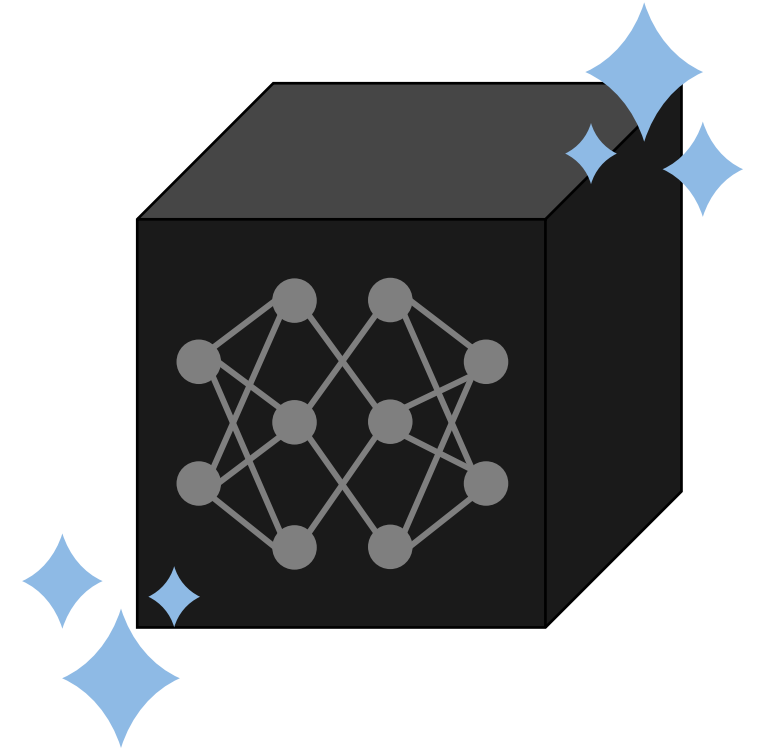


Large Language Models

You are here!

Das neue Werkzeug der Informatik : LLMs

- **Beliebiger** Text als Input, Text als Output
 - Multimodalität, Ton, Bild
 - Sehr **Intuitiv** zu verwenden
- **Black Box** (Aufgrund der Verwendung Neuronaler Netze)
 - LLM-Ausgaben sind **nicht vorhersehbar**
 - LLM-Ausgaben sind ggf. **nicht reproduzierbar**
 - LLM-Verhalten ist **schwer generalisierbar**
- **Hardware**
 - LLMs benötigen Unmengen an mehr, um zu laufen
 - Noch mehr um trainiert zu werden
 - Laufende Service-Kosten



Viele Vorteile durch Flexibilität und Transformative Eigenschaften



Home News Sport Business Innovation Culture Arts Travel Earth Audio Video Live

Airline held liable for its chatbot giving passenger bad advice - what this means for travellers

23 February 2024

Share Save

Maria Yagoda
Features correspondent

<https://www.bbc.com/travel/article/20240222-air-canada-chatbot-misinformation-what-travellers-should-know>

News

NYC's AI Chatbot Tells Businesses to Break the Law

The Microsoft-powered bot says bosses can take workers' tips and that landlords can discriminate based on source of income

By **Colin Lecher**

March 29, 2024 06:00 ET

<https://themarkup.org/news/2024/03/29/nycs-ai-chatbot-tells-businesses-to-break-the-law>

USA: Anwalt fällt auf von ChatGPT erfundene Urteile herein

Der Versuch eines Anwalts in New York, den Chatbot ChatGPT bei der Recherche für einen Fall zu verwenden, ist auf spektakuläre Weise schiefgegangen. Ein von ihm eingereichter Antrag enthielt Verweise auf Fälle wie "Petersen gegen Iran Air" oder "Martinez gegen Delta Airlines", die frei erfunden waren. Dem Anwalt zufolge wurden die angeblichen Urteile und Aktenzeichen dazu von ChatGPT ausgegeben.

Anwalt verwendete ChatGPT vor Gericht

In dem Fall reichte ein Passagier Klage gegen die Fluggesellschaft Avianca ein, weil er von einem Servierwagen am Knie verletzt worden sei. Die Airline beantragte, die Klage abzuweisen. In einem Gegenantrag aus dem März verwies die Anwaltskanzlei des

<https://rsw.beck.de/aktuell/daily/meldung/detail/usa-chatgpt-erfand-halbes-dutzend-urteile-fuer-klageantrag>

...

LLMs können **keine** Logischen Schlüsse aufstellen aber kreative Lösungen liefern

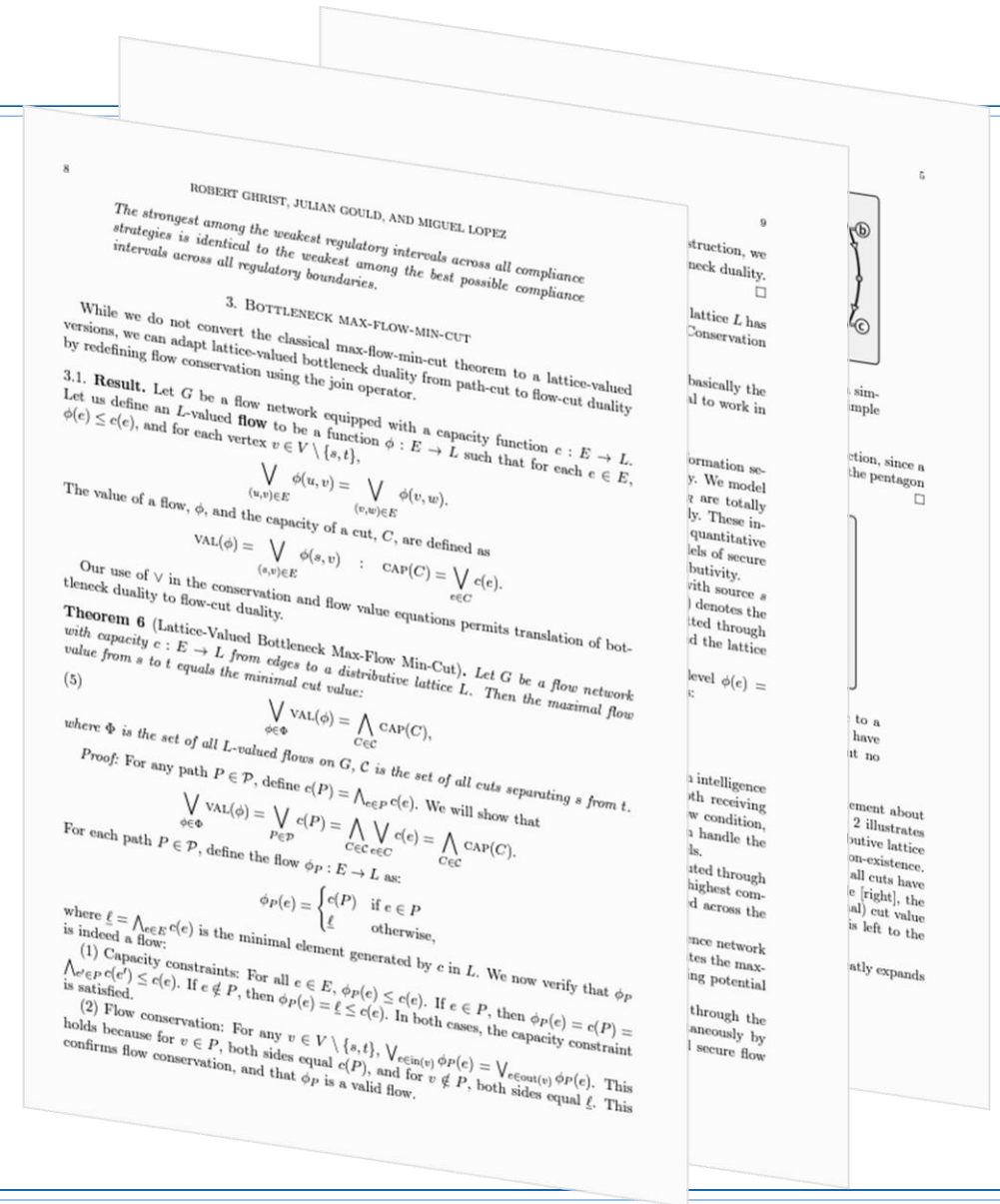
LATTICE-VALUED BOTTLENECK DUALITY

ROBERT GHRIST, JULIAN GOULD, AND MIGUEL LOPEZ

ABSTRACT. This note reformulates certain classical combinatorial duality theorems in the context of order lattices. For source-target networks, we generalize bottleneck path-cut and flow-cut duality results to edges with capacities in a distributive lattice. For posets, we generalize a bottleneck version of Dilworth's theorem, again weighted in a distributive lattice. These results are applicable to a wide array of non-numerical network flow problems, as shown. **All results, proofs, and applications were created in collaboration with AI language models.** An appendix documents their role and impact.

Eines der ersten Theoreme die mit einem LLM „bewiesen“ wurden.

Ghrist, R., Gould, J. and Lopez, M., 2024. Lattice-Valued Bottleneck Duality. *arXiv preprint arXiv:2410.00315*.



KI im Krankenhaus

Sana Krankenhäuser führen KI-Gestütztes Patient Data Management System (PDMS) ein

- KI als Assistenzsystem (Mensch-Maschine-Interaktion)
- Hohe Anzahl von Datenpunkten in der **Intensivmedizin**
- Aggregation und Priorisierung der Daten für Medizinisches Personal.
- Eingabe neuer Daten auf unkomplizierte Art und weise.
- Beschleunigung vieler Prozesse.



<https://www.aekno.de/aerzte/rheinisches-aerzteblatt/ausgabe/artikel/2024/november-2024/kuenstliche-intelligenz-als-chance-wahrnehmen>

Aktuell wird KI als Innovation in sehr vielen Software Produkten eingesetzt.

- Es gibt viele gute Anwendungsfälle für LLMs
- Es gibt aber auch viele Anwendungsfälle, die an Halluzinationen scheitern



Wo kommen Halluzinationen her ?
Gibt es Workarounds ?

Künstliche Intelligenz und Natürliche Dummheit



Linkshänder

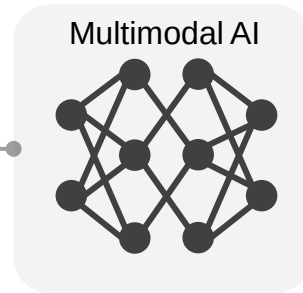


'Ferrari F80' als 5-Türer



Uhren die 8:00 Zeigen

Künstliche Intelligenz und Natürliche Dummheit



Von KI gezeichneter „Linkshänder“

Wie viele “E” sind in
dem Wort “ERDBEERE”

Nezhurina, M., Cipolina-Kun, L., Cherti, M. and Jitsev, J., 2024. Alice in Wonderland: Simple Tasks Showing Complete Reasoning Breakdown in State-Of-the-Art Large Language Models. *arXiv preprint arXiv:2406.02061*.

Wie viele “E” sind in
dem Wort “ERDBEERE”

4

Nezhurina, M., Cipolina-Kun, L., Cherti, M. and Jitsev, J., 2024. Alice in Wonderland: Simple Tasks Showing Complete Reasoning Breakdown in State-Of-the-Art Large Language Models. *arXiv preprint arXiv:2406.02061*.

Wie viele “E” sind in
dem Wort “ERDBEERE” 4

Alice hat 5 Schwestern und 3 Brüder.
Wie viele Schwestern hat ein
Bruder von Alice ?

Nezhurina, M., Cipolina-Kun, L., Cherti, M. and Jitsev, J., 2024. Alice in Wonderland: Simple Tasks Showing Complete Reasoning Breakdown in State-Of-the-Art Large Language Models. *arXiv preprint arXiv:2406.02061*.

Wie viele “E” sind in
dem Wort “ERDBEERE” 4

Alice hat 5 Schwestern und 3 Brüder.
Wie viele Schwestern hat ein
Bruder von Alice ?

6

Nezhurina, M., Cipolina-Kun, L., Cherti, M. and Jitsev, J., 2024. Alice in Wonderland: Simple Tasks Showing Complete Reasoning Breakdown in State-Of-the-Art Large Language Models. *arXiv preprint arXiv:2406.02061*.

Alice In Wonderland [3]

Alice hat 5 Schwestern und 3 Brüder.
Wie viele Schwestern hat ein Bruder von Alice ?

6

Wie viele “e” sind in dem Wort “Erdbeere”

4

LLMs können sehr gut Sprache imitieren, haben selbst
aber **kein Konzept von Logik**.

Mit technischen Erweiterungen kann man LLMs
aufwerten:

- Code Interpreter
- Plotter
- Function calls
- API integrationen

Nezhurina, M., Cipolina-Kun, L., Cherti, M. and Jitsev, J., 2024. Alice in Wonderland: Simple Tasks Showing Complete Reasoning Breakdown in State-Of-the-Art Large Language Models. *arXiv preprint arXiv:2406.02061*.



Daten : [3]



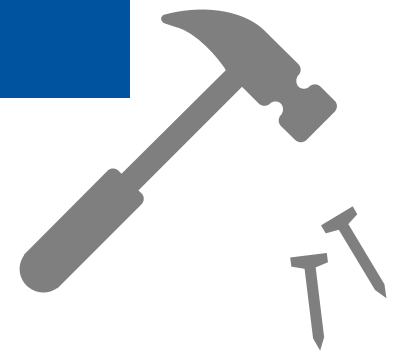
Künstliche Intelligenz (LLMs) ist nicht 'Intelligent' – aber dennoch extrem Hilfreich



LLMs ist ein weiteres algorithmisches Werkzeug das uns in der Software Entwicklung, Lehre und Forschung zur Verfügung steht.



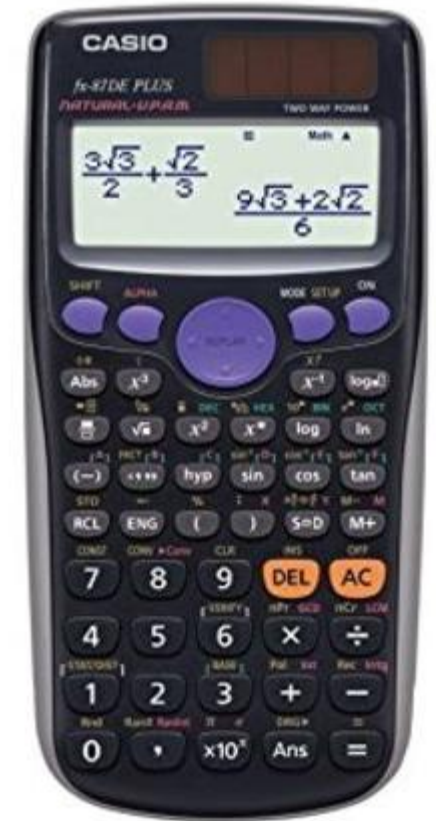
Wann sollte man LLMs einsetzen ?



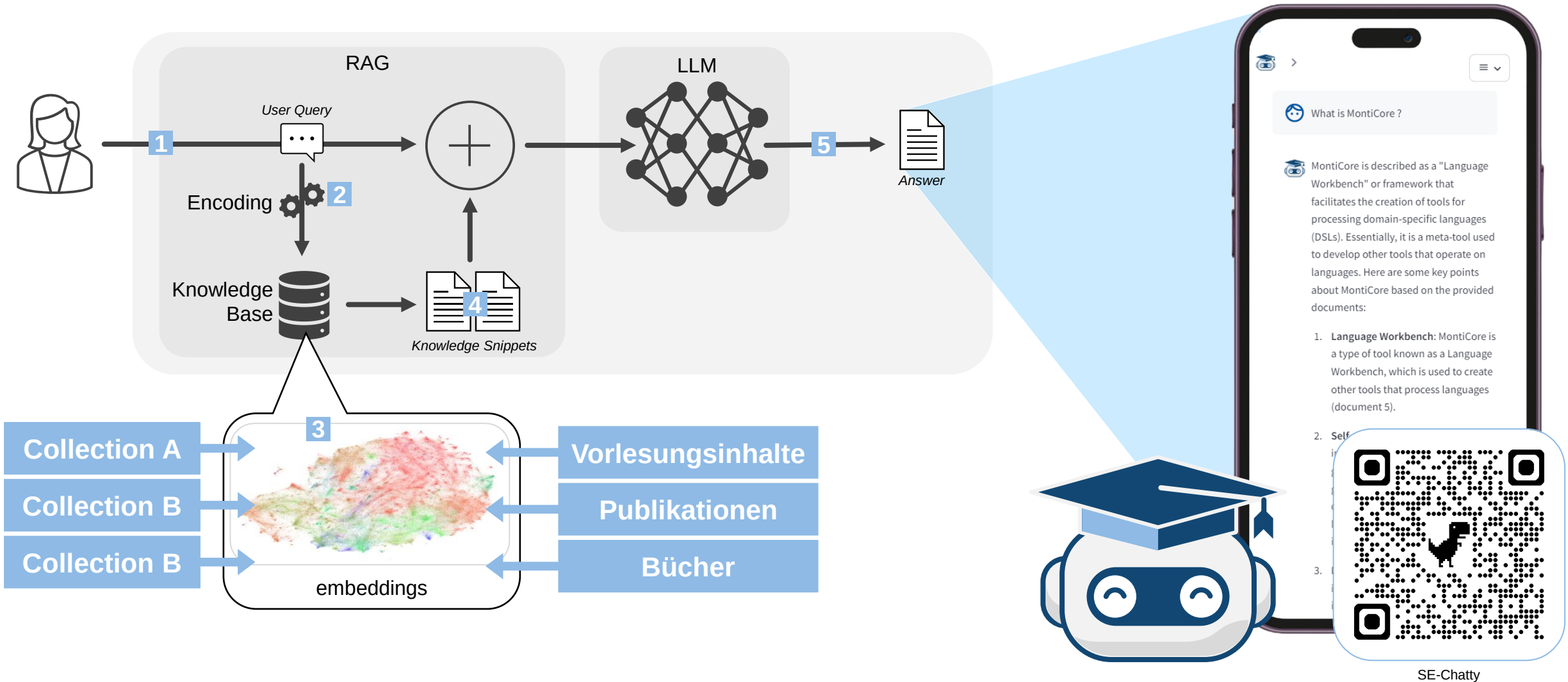
Verifizierbarkeit von Antworten

- LLMs als **Assistenz** nicht als 'Single Point of Truth'
- In Prozessketten: Antworten algorithmisch verifizieren.
- Anwender muss die **Kompetenz** besitzen Fehler zu identifizieren bzw. Antworten zu verifizieren.

$23 \times 9 = 301.230$?



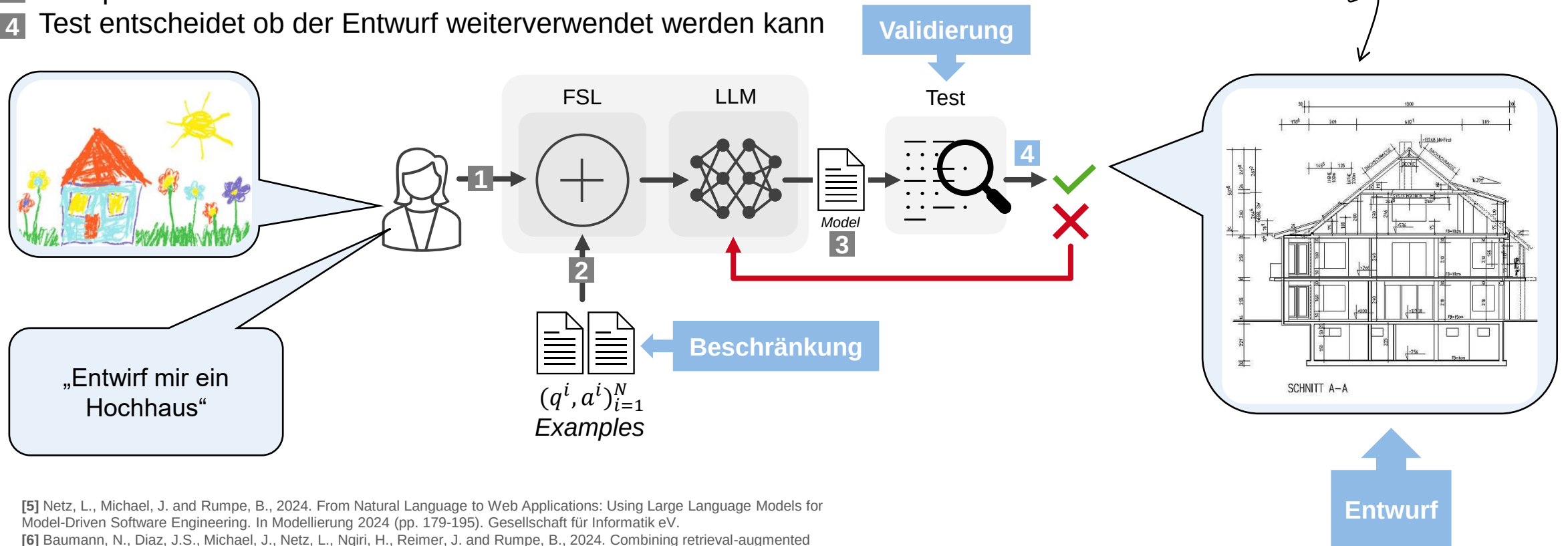
KI in der Lehre & Forschung: SE-Chatty - Integration von Internem Wissen & Know How



SE-Chatty

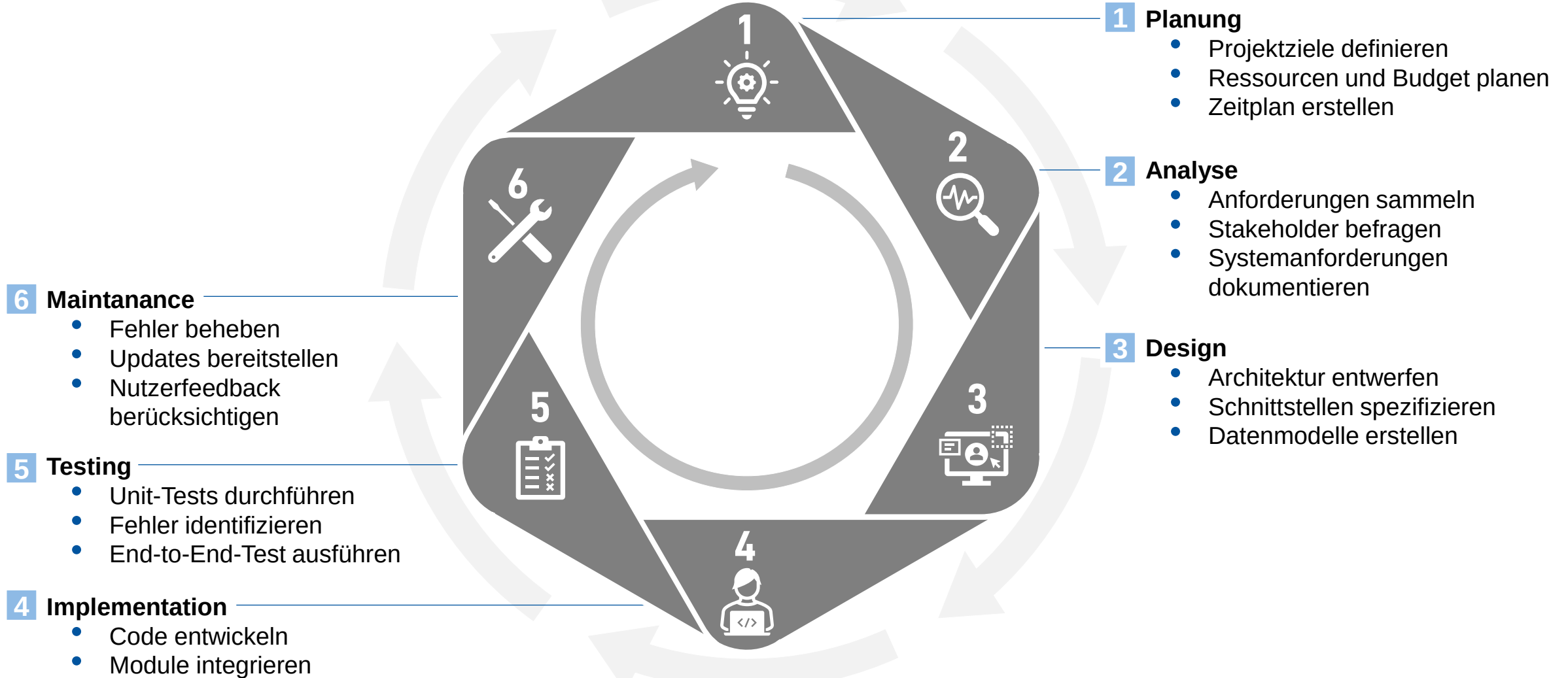
KI in SE: Kombination von Unterspezifikation und Formalisierung

- 1 Anwender kann in Freitext seinen Anwendungsfall beschreiben.
- 2 Prompt wird mit Formalismus erweitert.
- 3 LLM produziert Entwurf
- 4 Test entscheidet ob der Entwurf weiterverwendet werden kann



[5] Netz, L., Michael, J. and Rumpe, B., 2024. From Natural Language to Web Applications: Using Large Language Models for Model-Driven Software Engineering. In Modellierung 2024 (pp. 179-195). Gesellschaft für Informatik eV.

[6] Baumann, N., Diaz, J.S., Michael, J., Netz, L., Nqiri, H., Reimer, J. and Rumpe, B., 2024. Combining retrieval-augmented generation and few-shot learning for model synthesis of uncommon DSLs. In Modellierung 2024 Satellite Events (pp. 10-18420). Gesellschaft für Informatik eV.



KI in der Software Entwicklung

- KI ist in der Software Entwicklung mittlerweile weit verbreitet.
- KI soll Produktivität um 35% prozent steigern
- Salesforce stellt keine Programmierer mehr ein.
- 25% des neuen Sourcecodes von Google ist KI-Generiert [1]

[1] <https://blog.google/inside-google/message-ceo/alphabet-earnings-q3-2024/#full-stack-approach>

Career / News

Salesforce Will Hire No More Software Engineers in 2025, Says Marc Benioff



By Henry Martin
December 18, 2024

Salesforce will not be hiring any more software engineers in 2025 amid significant productivity boosts from AI, Marc Benioff has revealed.

<https://www.salesforceben.com/salesforce-will-hire-no-more-software-engineers-in-2025-says-marc-benioff/>

Navigating the generative AI disruption in software

June 5, 2024 | Article

Code development through gen AI is another significant source of value creation for software companies. Recent estimates indicate the technology can **improve developer productivity by 35 to 45 percent**, a spike that outperforms past advances in engineering productivity, leading to lower cost of code development. Gen AI can also speed up the processes of documenting code

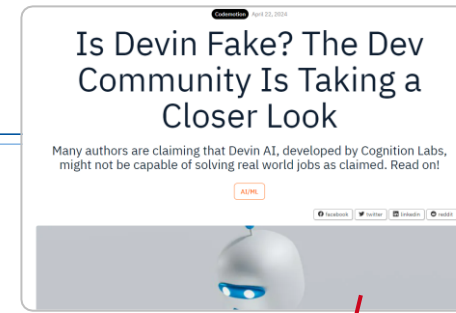
<https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/navigating-the-generative-ai-disruption-in-software>

KI-Hype // BLASE

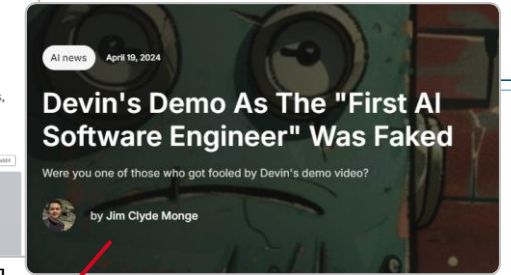
- Unsicherheit bei Framework / Modell-wahl
- Wenig Erfahrungswerte mit bestehenden Technologien
- Starke Fluktuation in den Angebotenen Software- und Hardwarelösungen im Bereich KI
- Unklarer Stand der Technik
- ABER: Hype „flaut“ gerade ab



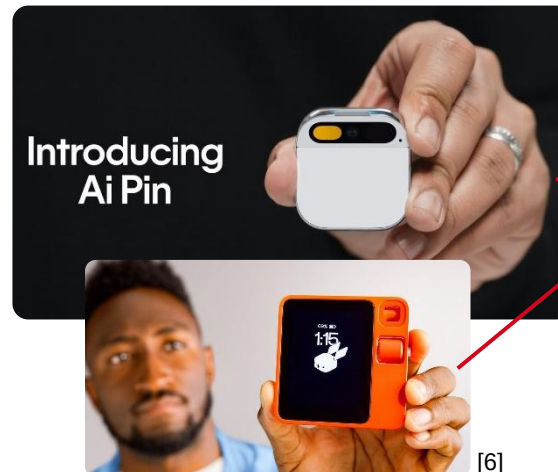
[3]



[4]

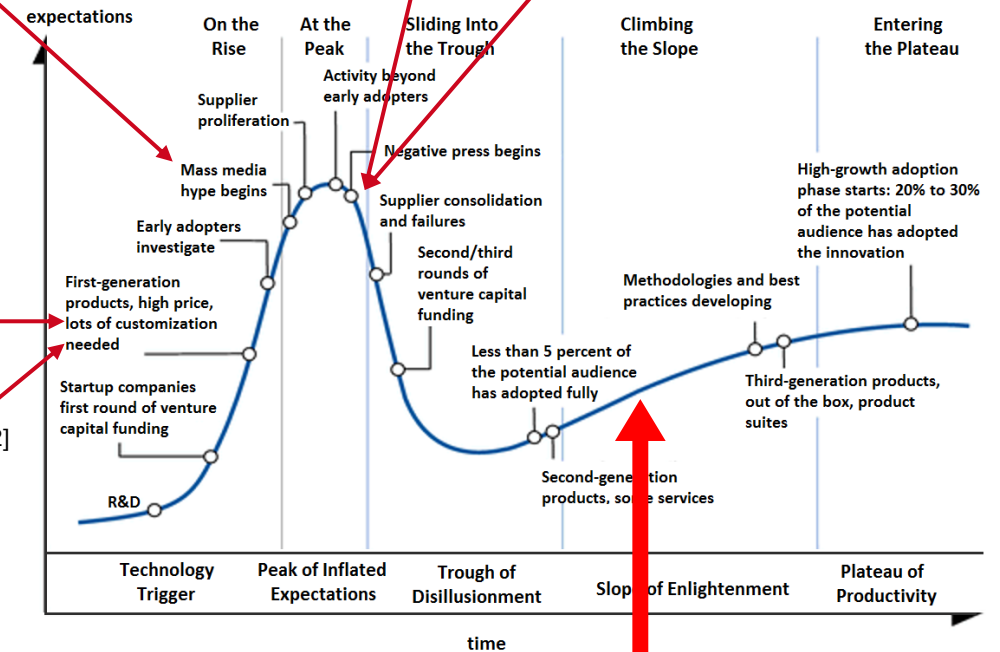


[5]



[2]

[6]



General Hype Cycle for Technology

You are here!

[2] <https://humane.com/aipin>

[3] <https://www.geeksforgeeks.org/devin-ai-worlds-first-ai-software-engineer/>

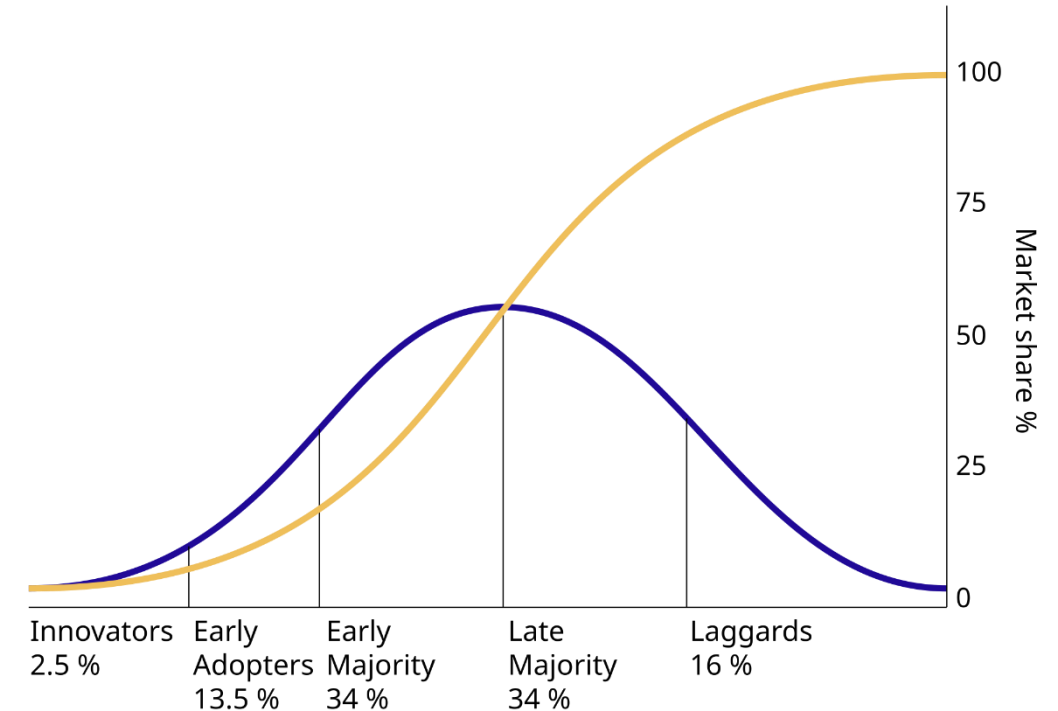
[4] <https://www.codemotion.com/magazine/ai-ml/is-devin-fake/>

[5] <https://www.zeniteq.com/blog/devins-demo-as-the-first-ai-software-engineer-was-faked>

[6] Rabbit R1: Barely Reviewable <https://www.youtube.com/watch?v=ddTV12hErTc>

Produktadoption

- **Innovatoren 2,5%** (OpenAI, StartUps,)
Risikofreudige und technikaffine Unternehmen, die neue Produkte als Erste ausprobieren.
- **Frühe Unternehmer 13,5%** (Microsoft Copilot, Google ...)
Meinungsführer, die Neuerungen schnell annehmen und deren Erfahrungen andere beeinflussen.
- **Frühe Mehrheit 34%** (Innovative KMUs)
Unternehmen, die Neuerungen akzeptieren, sobald sie sich im Markt bewährt haben.
- **Späte Mehrheit 34%** (Großbetriebe)
Skeptische Konsumenten, die erst nach der breiten Akzeptanz eines Produkts folgen.
- **Nachzügler 16%** (Ämter, Traditionsunternehmen...)
Traditionsbewusste Personen, die Veränderungen meiden und Innovationen erst spät übernehmen.



Zusammenfassend:



Wann sollte man LLMs einsetzen ?

- In nicht-kritischen oder in Verifizierbaren Kontexten
- In transformativem Kontext um informelle/ unstrukturierte Daten zu Formalisieren und maschinell weiter zu verarbeiten
- Nicht um Logische Schlüsse durch eine Maschine zu fassen.
- Wenn eine Algorithmische alternative fehlt / nicht praktikabel ist.



Wie können LLMs im SE, Forschung und Lehre helfen ?

- Können verifizierbar Wissen bereitstellen sowohl für Studenten als auch für Wissenschaftler bereitstellen
- Können in Explorativen Aufgaben unterstützen.
- Können bei der Code Erstellung und im SDLC assistieren

Vielen Dank für Ihre Aufmerksamkeit
Gibt es Fragen ?