

KI zum Frühstück

Generative KI-Systeme

Was ist KI?

Technologien, die intelligentes Verhalten nachahmen und einen gewissen Grad an **Eigenständigkeit aufweisen, um bestimmte Aufgaben zu erledigen** (z. B. Machine Learning, Deep Learning, Text-, Sprach- oder Bilderkennung).

KI-Technologien umfassen sowohl rein software-basierte Systeme (z. B. Chatbots, Spracherkennung, Übersetzungen oder zur Datenanalyse basierend auf Machine Learning) **als auch software-basierte Systeme, die in autonome Geräte integriert sind** (z. B. autonome Drohnen oder Fahrzeuge, Roboter in Lagerverwaltungs- oder Produktionsprozessen).

Quelle: <https://www.statistik.at/fileadmin/announcement/2024/10/20241016IKTU2024.pdf>

AI vs. AGI

Von „echter KI“ können wir reden, wenn AGI* startet.

* (Artificial General Intelligence / Allgemeine Künstliche Intelligenz)



Spezialisierte Fähigkeiten



Allgemeine Fähigkeiten



Datenabhängiges Lernen



Selbstständiges Lernen



Mustererkennung



Konzeptionelles Verständnis



Menschliche Anleitung



Hohe Autonomie



Statistiken

Wo steht KI in den Umfragen?

- **Jede/r 3. Deutsche hat Angst vor den Folgen von KI / Jede/r 4. hat Sorgen um Datenmissbrauch im Netz.**

(Quelle: 19.11.25: <https://www.ruv.de/newsroom/themenspezial-die-aengste-der-deutschen/pressemitteilungen/2025-11-19-angst-folgen-ki>)

- **50% der Amerikaner sind eher besorgt als begeistert über zunehmenden Einsatz von KI im Alltag.**

(Quelle: PEW Research Center – 17.09.2025:
<https://www.pewresearch.org/science/2025/09/17/how-americans-view-ai-and-its-impact-on-people-and-society/>)

- **China (83%), Indonesia (80%) & Thailand (77%), sehen mehr Vorteile in KI als Gefahren.**

(Quelle: Stanford HAI 2025: <https://hai.stanford.edu/ai-index/2025-ai-index-report>)

Der EU AI Act



Unzulässige KI-Anwendungen

KI-Systeme, die als besonders riskant eingestuft werden und vollständig verboten sind.

zB: social scoring, predictive policing, biometrische Erfassung/Überwachung...



Hochrisiko-KI-Anwendungen

KI-Systeme, die strengen Anforderungen unterliegen, um Sicherheit und Fairness zu gewährleisten.

zB: Aufzüge, Spielzeug, Agrar & Forstfahrzeuge, Medizinische Geräte...



Geringes Risiko-KI-Anwendungen

KI-Systeme mit geringem Risiko, die grundlegende Sicherheitsstandards erfüllen müssen.

zB.: Chatbots, Text/Audio/Video/Bild generative KI



Minimalrisiko-KI-Anwendungen

KI-Systeme, die kaum reguliert sind und kein signifikantes Risiko darstellen.

zB.: Computerspiele, Spamfilter

Transparenz

- Userinnen informieren dass sie mit einem KI-System interagieren und vor Fehlinformationen warnen.
- KI-Inhalte müssen auf Social Media und Website gekennzeichnet werden.

1. VOR DER INTERAKTION INFORMIEREN



2. KI-INHALTE KENNZEICHNEN



EU AI Act Quellen

EU Seite: <https://artificialintelligenceact.eu>

RTR Seite: https://www.rtr.at/rtr/service/ki-servicestelle/ai-act/risikostufen_ki-systeme.en.html

Wichtig:

Es werden immer wieder Anpassungen und Details hinzugefügt. Es zählt sich also aus immer wieder in den Quellen nachzulesen!

BIAS & Halluzination

& generelle Stolperfallen

Stolperfallen



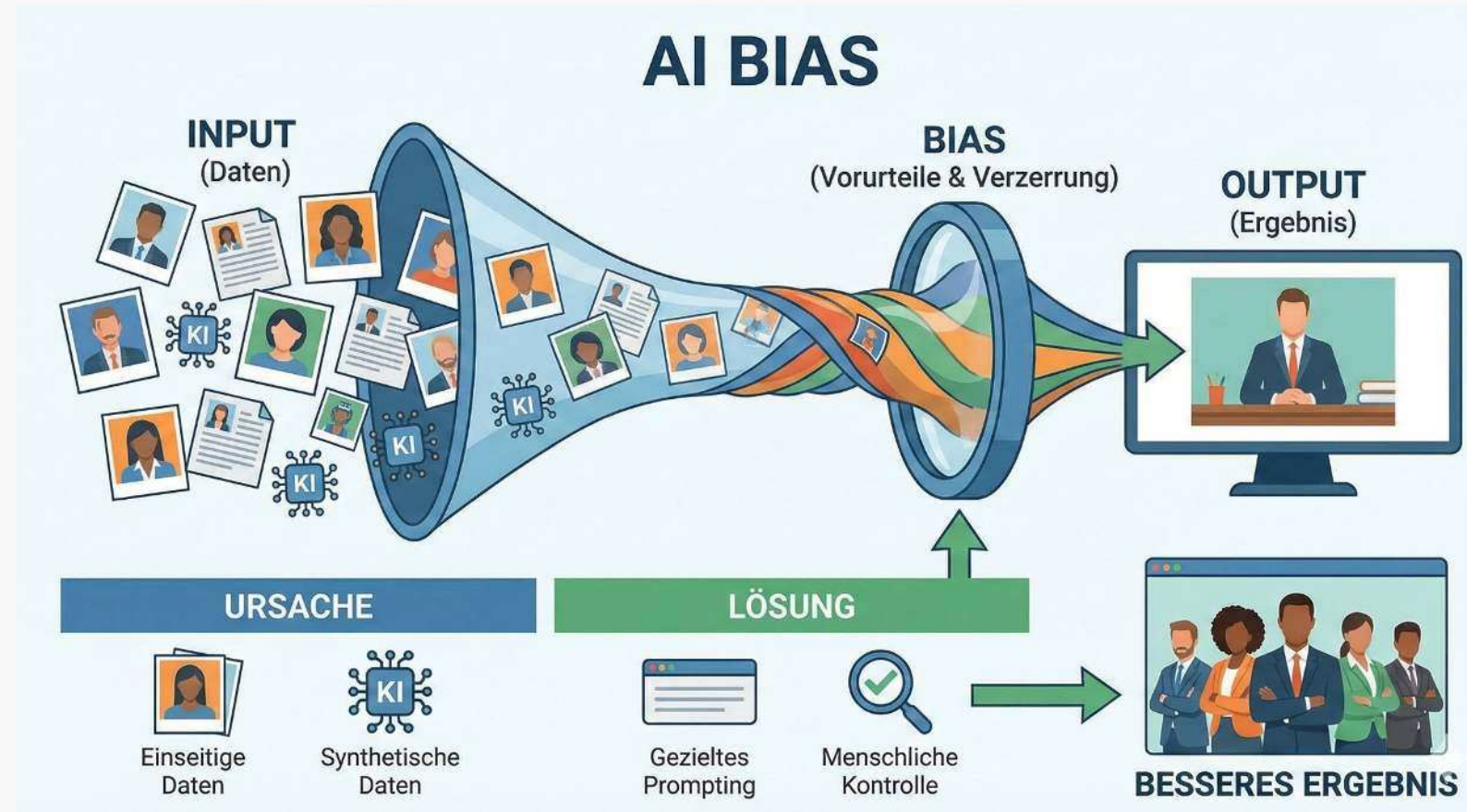
BIAS

BIAS?

Voreingenommenheit, bei der verzerrte Daten das Gesamtbild der KI-Ausgabe verfälschen.

Gegenmaßnahmen:

- **Gezieltes Prompting:** präzise Anweisungen können einseitige Muster umgehen.
- **Datenprüfung:** Ergebnisse und Quellen müssen kritisch auf ihre Objektivität hinterfragt werden.



BIAS Studien

Diverse Studien befassten sich mit BIAS Themen in der Bilderstellung. Hier eine Auswahl:

- <https://arxiv.org/pdf/2409.19959>
- <https://www.mdpi.com/2313-433X/11/2/35>
- <https://arxiv.org/html/2503.23398v1#S1>

Hier ist klar zu erkennen, dass gezieltes Prompting und Wachsamkeit der Userinnen eine entscheidende Rolle bei der Vermeidung von BIAS-Inhalten spielen.

Typische Anwendungsfälle

Gängige Anwendungen

E-Mails

Vertrieb

Recruiting

Übersetzungen

Gängige Anwendungen II

Bildgenerierung

Content Creation

(Social Media Beiträge, Blog,
Produkt-beschreibungen...)

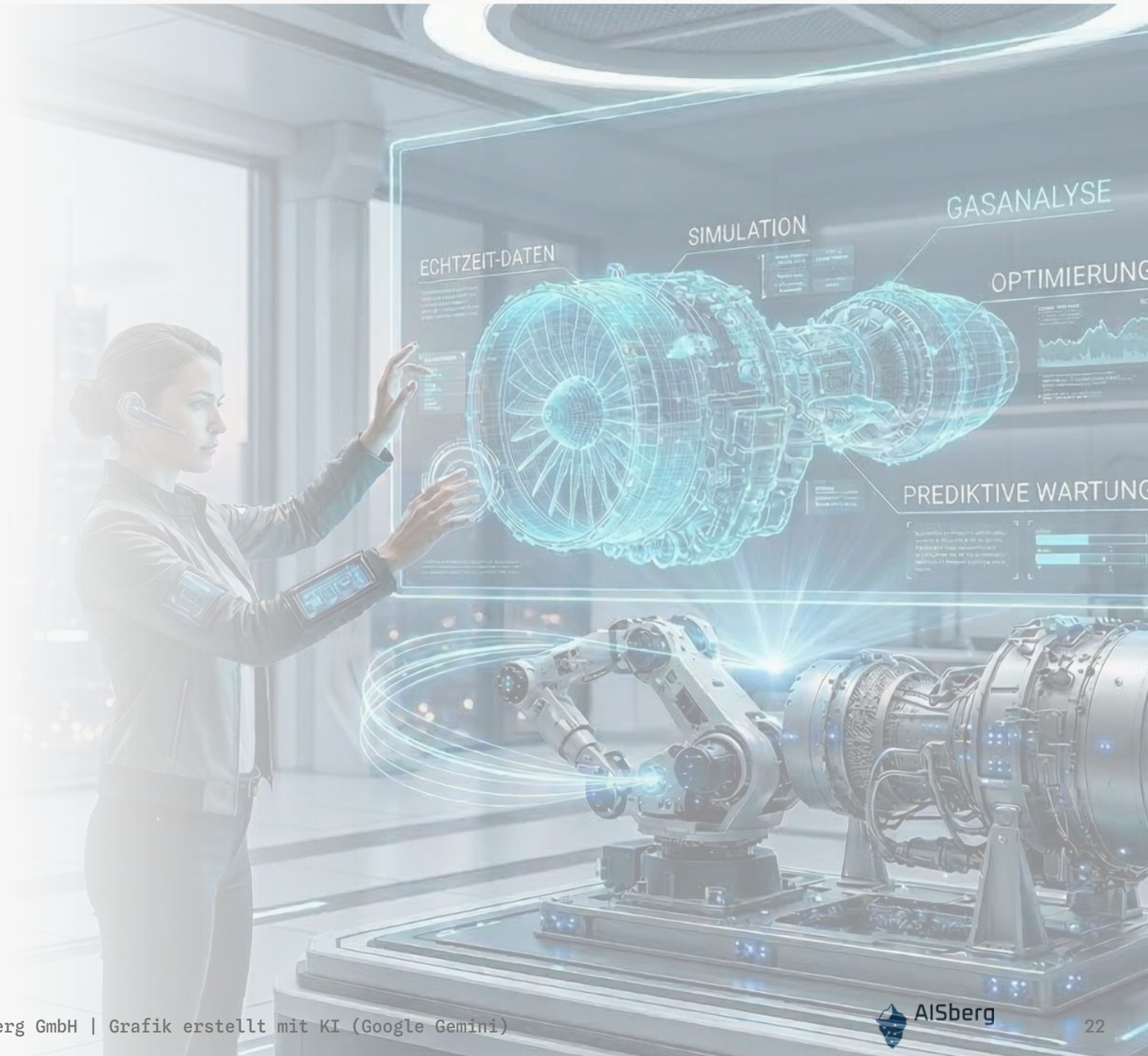
Zusammenfassungen

Digitale Zwillinge

(Personen & Industrie)

Digitale Zwillinge

Industrie
Warehouse
Logistik
Personen



Digitale Zwillinge

- **Digitale Zwillinge**

CEOs lassen digitale Zwillinge von sich erstellen

(Quelle: <https://www.deutschlandfunk.de/ki-chef-digitale-zwillinge-100.html>)

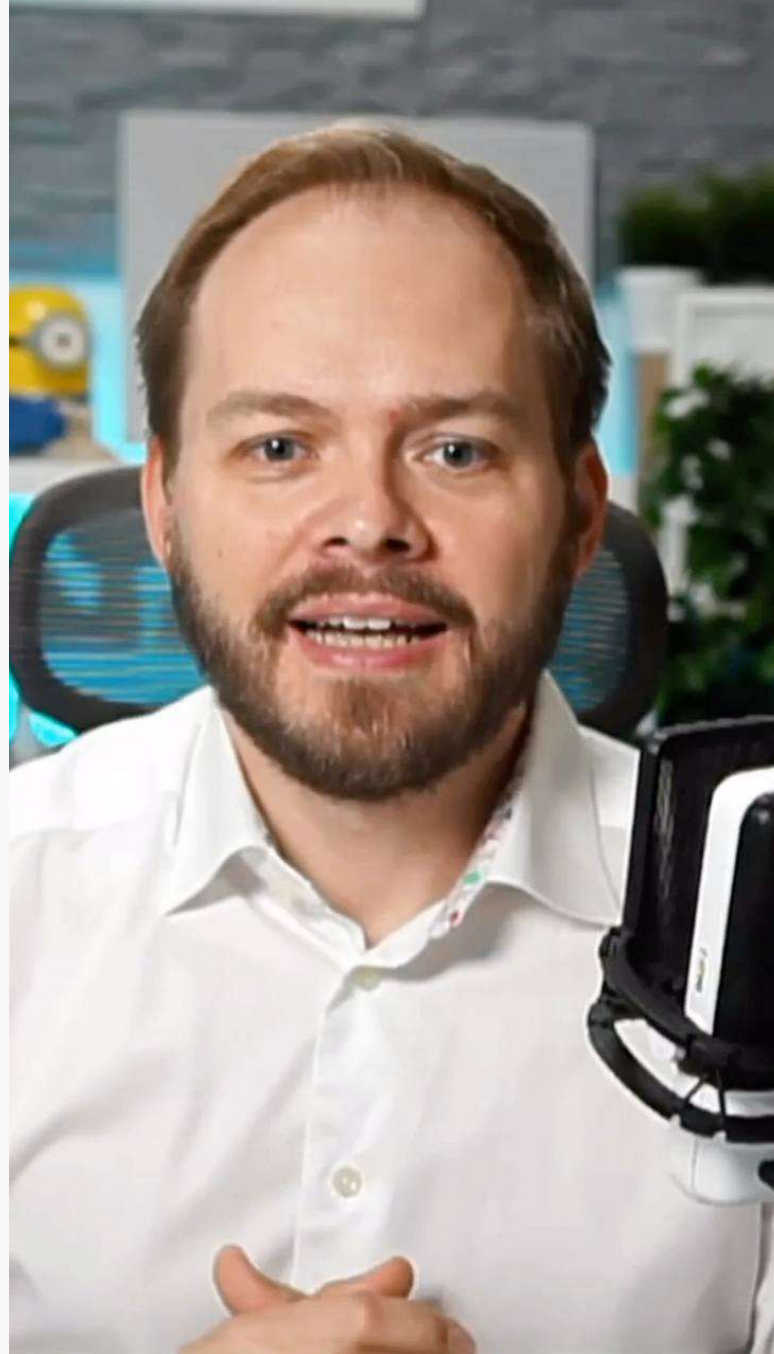
- **H&M erstellt DZ von Models**

(Quelle: <https://www.bbc.com/news/articles/c3vvg73xndeo>)

Stimmklonung



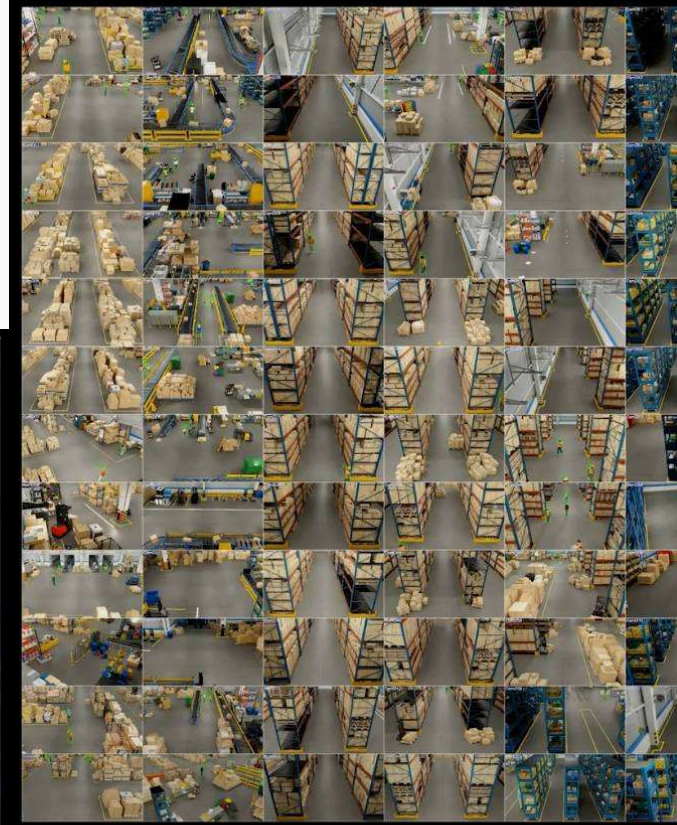
(Video) Avatar Beispiel



Lager

- CAT Zeichnungen in Nvidia Omniverse
- Virtuelle Kameras (Metropolis von Nvidia)
- Wegpunkte mit (CuOpt)
- Isaac-Plattform

<https://www.youtube.com/watch?v=15M4sqaRd6w>



POS - Zwillinge

Wie können Kassensysteme positioniert werden?

Bevor Investitionen getätigt werden.

<https://www.youtube.com/watch?v=Q1SBBtAPcVk>



IKEA

*CGI ist bei IKEA längst normal

- Schneller
- Flexibler
- Sicherer (Feuer/Wasser)
- AR (Möbel in Wohnung platzieren)

Quelle:
<https://blog.chaos.com/putting-the-cgi-in-ikea-how-v-ray-helps-visualize-perfect-homes>



Bildquelle: <https://blog.chaos.com/putting-the-cgi-in-ikea-how-v-ray-helps-visualize-perfect-homes>

*CGI = Computer Generated Imagery

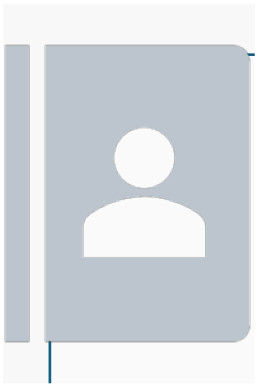


Verantwortung

Input & Output

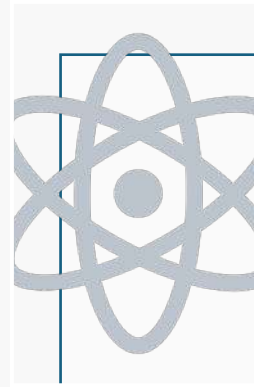
Die Verantwortung, bei der Nutzung einer öffentlichen KI, um DSGVO- & Cybersecurity-Verletzungen zu verhindern, liegt bei allen Nutzerinnen!

Eingabe-Verantwortung



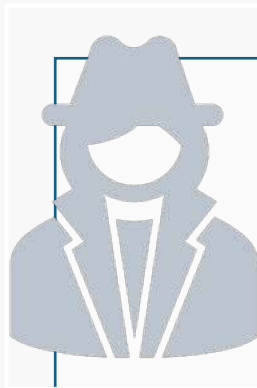
Keine personenbezogenen Daten:

Namen, E-Mail-Adressen, Telefonnummern, Kundennummern oder Personalakten, usw. dürfen nicht eingegeben werden.



Keine Geschäftsgeheimnisse:

Unternehmensinterna, Finanzdaten, Strategiepapiere, unveröffentlichte Berichte oder Vertragsinhalte sind tabu.



Anonymisierung & Pseudonymisierung

Wenn Daten genutzt werden, müssen sie vollständig anonymisiert sein. Besser ist es, allgemeine Informationen zu nutzen.

Klare Verantwortung

Im Rahmen der AI Governance sollte im Unternehmen klar und einfach kommuniziert werden, welche Informationen von den Nutzerinnen in einer öffentlich zugängigen KI verarbeitet werden dürfen.

Hierzu zählt auch die Nutzung KI-Unterstützter Tools & Systeme

Verantwortung: Output

Kennzeichnungspflicht, Urheberrecht



OUTPUT REVIEW

HUMAN
OVERSIGHT
ACTIVE

Mensch hat weiterhin Verantwortung

- Userinnen sollten den Output einer KI nicht ungeprüft veröffentlichen oder freigeben.
- Eine gewisse Fachexpertise, zum abgefragten Thema ist weiterhin vorteilhaft, um Ergebnisse zu prüfen.
- Personen, Zahlen, Datumsangaben, Zitate, historische Fakten, Quellen sind häufige Stolperfallen.
- Achtung auch bei der Bildgenerierung! Logos können wir schnell in eine Copyright-Infringement-Falle führen.

Logo-Falle

Midjourney

Explore

Create

Edit

Personalize

Moodboards New!

Organize

Chat

Tasks

Help

Updates

Dark Mode

philippbell...

What will you imagine?



riki_64182 1d

<character reference link> A magical Pixar-Disney style illustration of the same two elegant young women on a Paris rooftop terrace at sunset, celebrating with sparkling champagne. The blonde woman laughs while holding a glass, the brunette raises hers in a toast. The skyline glows in golden light, the Eiffel Tower visible in the distance, soft pink clouds painting the horizon. The mood is cinematic and dreamy, with gentle wind moving their hair, bokeh lights in the background, and warm sunlight on their faces. Their expressions are natural, vivid, full of emotion, with luminous eyes and perfect Pixar-style detailing. A story of friendship, joy, and elegance above the rooftops of Paris.

ar 9:16

style raw

v 7

stylize 900

Use

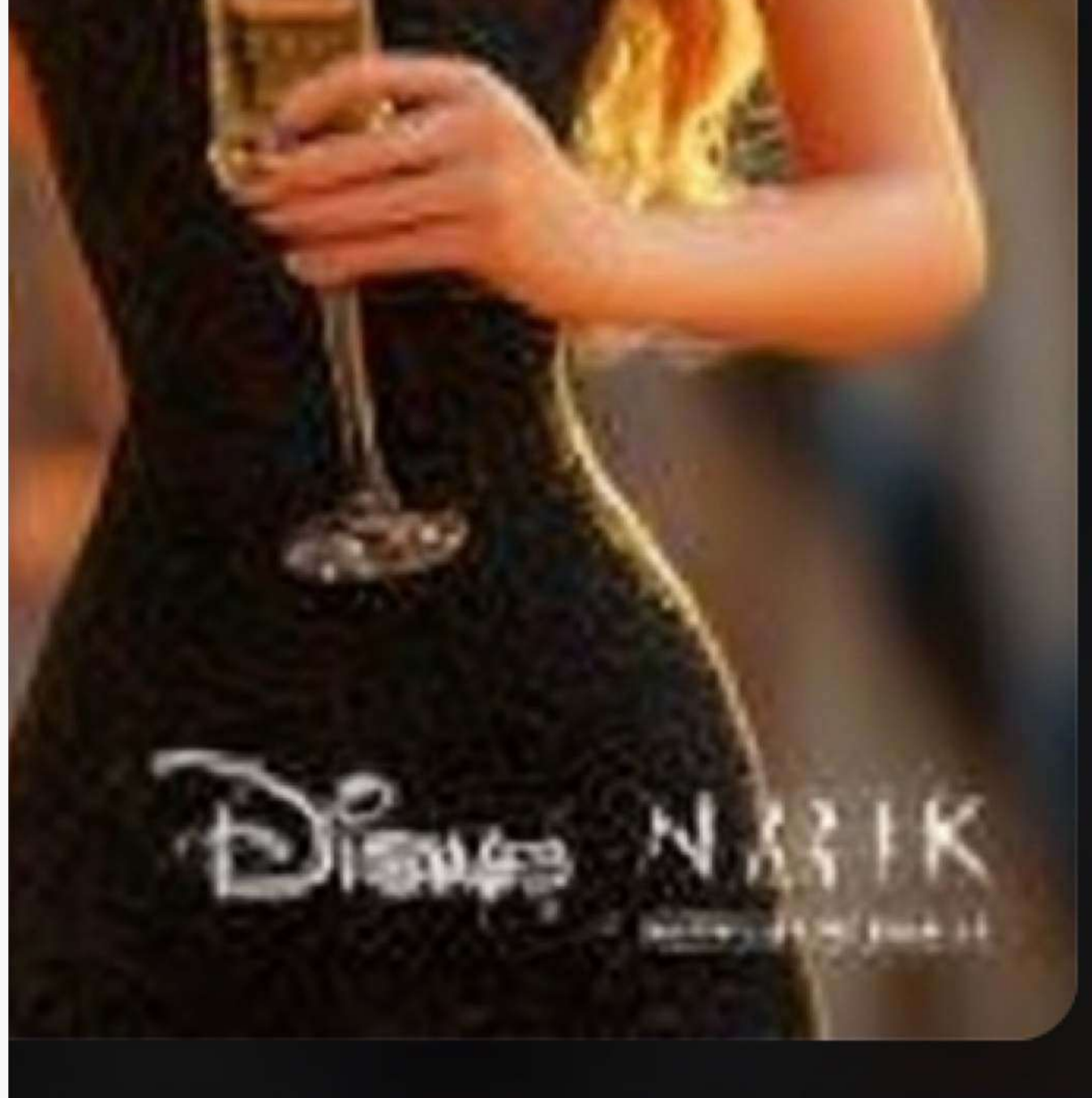
Image

Style

Prompt

Logo-Falle

Achtung bei solchen Bildern: Logo von Disney ist vielleicht nicht 100% zu erkennen, jedoch gut genug um zu wissen dass es sich um Disney handelt!



Beispiel – Werbung Flughafen Frankfurt



Beispiel – Werbung Flughafen Frankfurt

GER
58

and
Filippo's.

 Frankfurt
Airport

**GOODMAN
& FILIPPO**



350 m

Risiko: Cloud

Aktuelle Themen

Vorteile/Nachteile

Gefahren

Daten(schutz) & Big-Tech

- **starke De-Regulierung in USA - Juli 2025 - Winning the Race - America's AI Action Plan**
(Quelle: <https://www.whitehouse.gov/articles/2025/07/white-house-unveils-americas-ai-action-plan/>)
- **Microsoft sperrt Dienste des IStGH Chefklägers Karim Khan**
(Quelle: <https://www.heise.de/news/Strafgerichtshof-Microsofts-E-Mail-Sperre-als-Weckruf-fuer-digitale-Souveraenitaet-10387368.html>)
- **Microsoft gibt zu: EU-Daten sind nicht sicher vor einer Übermittlung an die USA**
(<https://www.heise.de/news/Nicht-souveraen-Microsoft-kann-Sicherheit-von-EU-Daten-nicht-garantieren-10494684.html>)
- **Palantir in Bayern (Hessen & Nordrhein-Westfalen)**
(Quelle: <https://www.br.de/nachrichten/deutschland-welt/polizeisoftware-von-palantir-werden-alternativen-ignoriert,UzmaB7J>)
- **OpenAI wurde von US-Bundesgericht beauftragt Chats zu speichern (aufgehoben am 09.10.2025 - US-Richterin Ona Wang)**
(Quelle: <https://www.golem.de/news/richterliche-anordnung-openai-muss-keinen-temporaeren-chatgpt-chats-mehr-speichern-2510-201091.html>)
- **China will globale KI-Regulierung Organisation**
(Quelle: <https://the-decoder.de/china-plant-neue-organisation-fuer-weltweite-ki-regulierung/>)

Open Source vs. Big-Tech

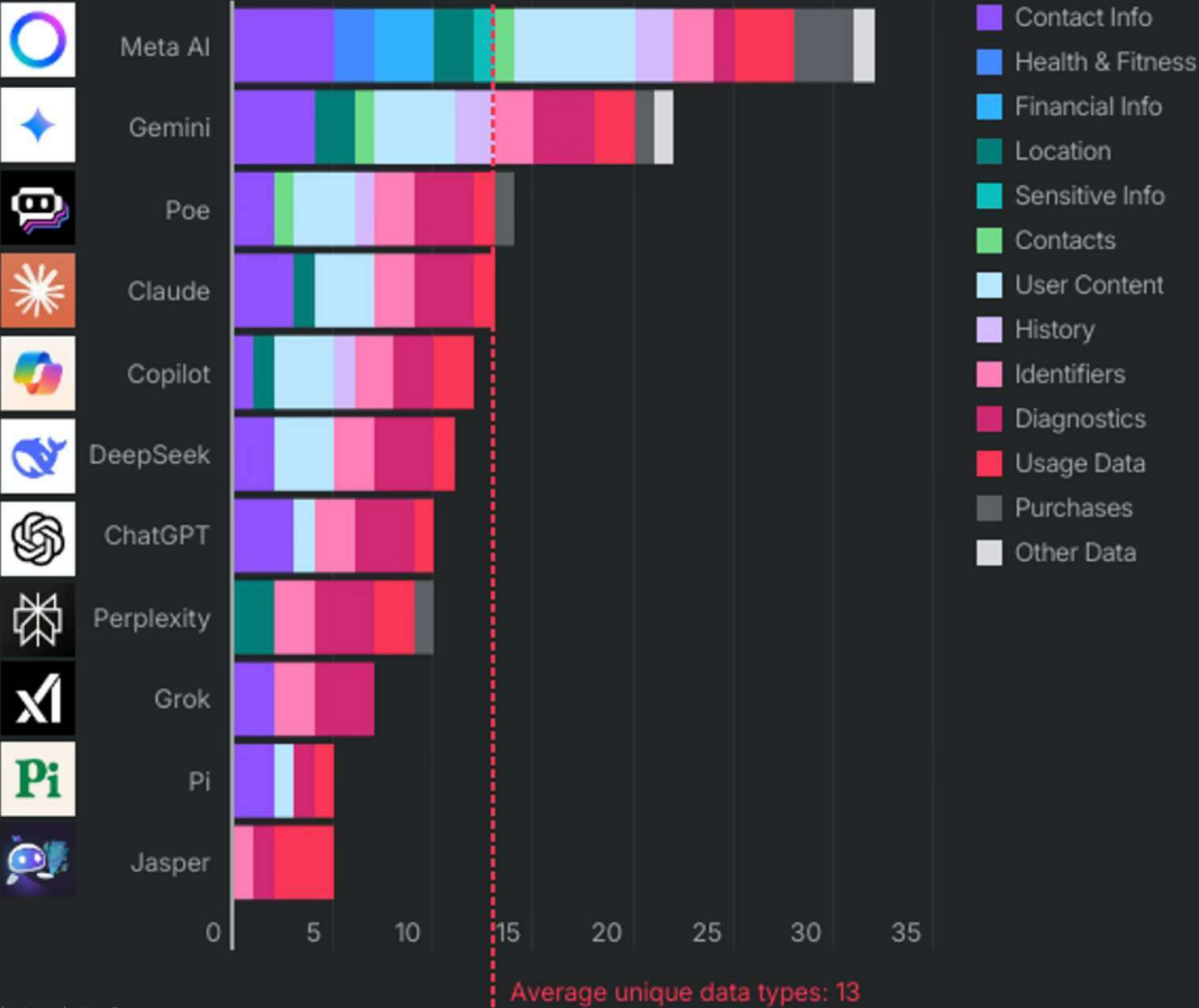
- **IStGH führt jetzt OpenDesk aus DE ein und verabschiedet sich von Microsoft.**
(Quelle: <https://www.handelsblatt.com/technik/it-internet/software-straengerichtshof-ersetzt-microsoft-durch-deutsche-loesung/100166382.html>)
- **Deutsche Forschungsgemeinschaft (DFG) zieht Forschungsdaten von US-Clouds zurück - aufgrund des Cloud Acts.**
(Quelle: <https://www.speicherguide.de/infrastrukturen/dfg-holt-forschungsdaten-aus-us-clouds-zurueck-26867.html>)
- **Schweizer Datenschützer empfehlen breites Cloud-Verbot (AWS, Google, Microsoft).**
(Quelle: <https://www.heise.de/news/Schweiz-Datenschuetzer-verhaengen-breites-Cloud-Verbot-fuer-Behoerden-11093438.html>)
- **Frankreich ersetzt Microsoft Teams und Zoom durch Visio.**
(Quelle: <https://www.heise.de/news/Frankreich-ersetzt-MS-Teams-und-Zoom-durch-eigene-Videokonferenzsoftware-11155120.html>)

Datenkrake - Cloud

DATA UPDATE: MAY 20, 2025

Meta AI collects over 90% of user data types

Among AI chatbots, Meta AI collects the most user data, with 32 out of 35 data types, which is more than double the average of 13.



https://surfshark.com/research/chart/ai-chatbots-privacy?srsltid=AfmB0opqIsj_3tpeeTNDJTLw4dipFPeJ2QoBRj1MvvrA2bbkYpWdrN_G

Cloud vs. Local

Cloud

Vorteile:

- Schnell & „einfach“
- Multi-Talent (Text/Video/Bild/Audio...)

Nachteile:

- Evtl. eingeschränkter Funktionsumfang
- nicht klar, was mit den Daten passiert.
- Werbung wird Ergebnisse beeinflussen – weiteres BIAS
- PII werden schnell und unbedacht eingegeben.

Local

Vorteile:

- Alle Daten liegen im Unternehmen
- Volle Kontrolle über Daten und KI-Modelle
- Model-Auswahl vielfältig (2,6+ Mio.)
- eher spezialisiert als Allrounder
- Daten verlassen das Unternehmen nicht

Nachteile:

- Systemanforderungen
- Model Auswahl treffen

Tools

Desktop

Ollama → <https://ollama.com/>

LM Studio → <https://lmstudio.ai>

- Einfach Installation
- RAG Möglichkeit direkt integriert
- Direkt Zugriff auf Huggingface
- Modelle können über Benutzeroberfläche direkt heruntergeladen werden

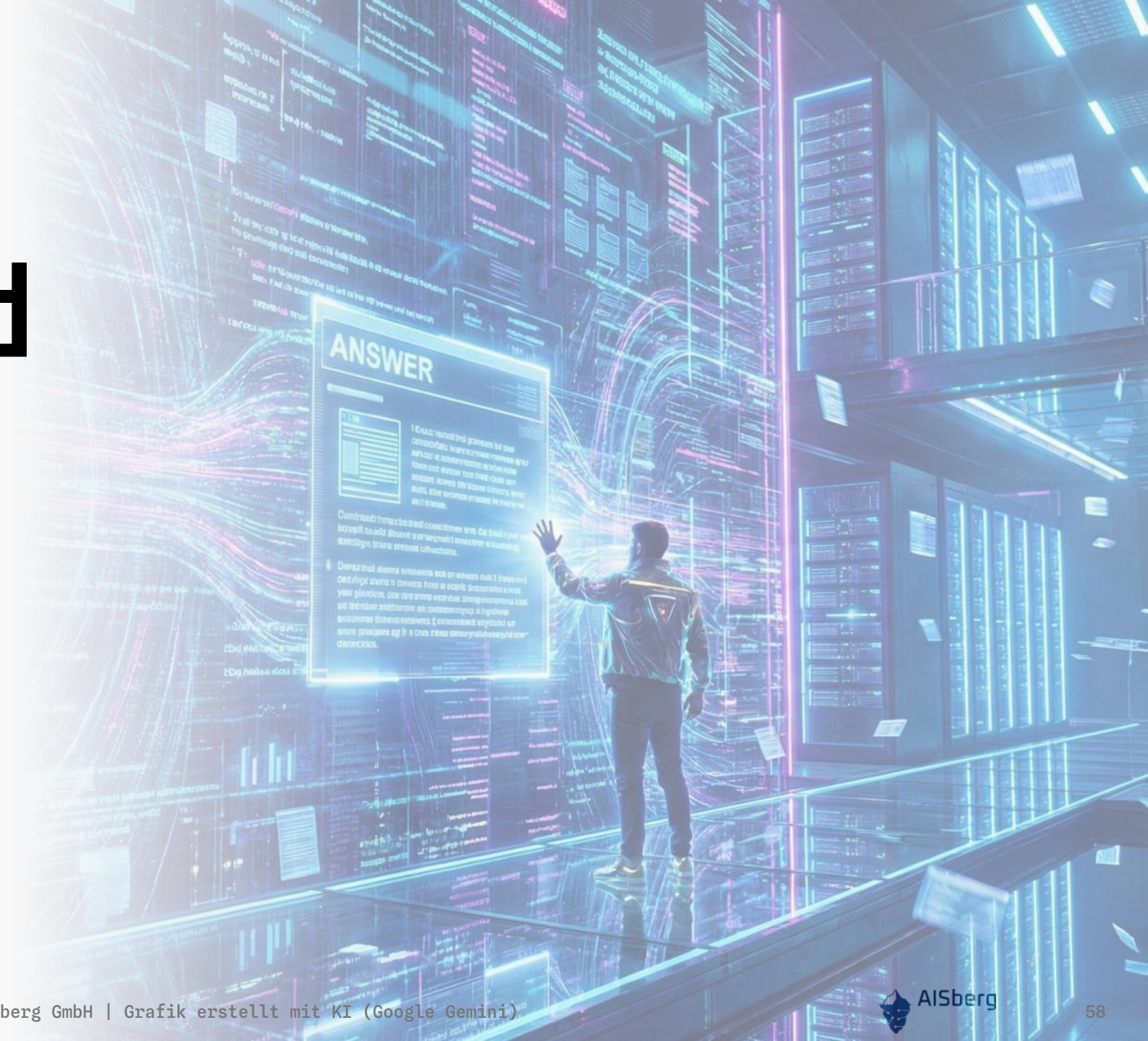
Mobile

- **Private LLM (iOS):**
<https://privatellm.app/de>
- **Private AI (Android):**
<https://play.google.com/store/apps/details?id=us.valkon.privateai>
- **Pocket Pal (Android & iOS):**
<https://github.com/a-ghorbani/pocketpal-ai> (links zu Stores etwas weiter in der Mitte) – Empfehlung

Lokale KI – Vorteil von SLM gegenüber LLM

- kleine **SLM** (Small Language Models) agieren als **Spezialisten** – im Gegensatz zu LLM (Large Language Models).
- Lokale KI gewährleistet **Datensouveränität** und **Sicherheit**
- **Compliance** Themen sind leichter einzuhalten
- Branchenspezifisches Wissen kann integriert werden (Fine-Tuning)
- Energieeffizienter
- Mehrere SLM können zusammenarbeiten
- Anbindung an RAG

Retrieval Augmented Generation



RAG

- Unternehmenswissen den Mitarbeitenden verfügbar machen.
- Mittels natürlicher Sprache können auch komplexere Anfragen ohne zwingendes spezifisches Wissen durchgeführt werden.
- Dokumente werden für KI Suche aufbereitet und von einem Embedding-Model verwaltet. Zugleich bleiben sie im klassischen Archiv vorhanden.
- KI bezieht Wissen aus RAG-Datenbank.
- Neues Wissen kann jederzeit ergänzt werden. Altes entfernt.

Livedemo

LIVEDEMO



NETIVE PFEUEKT:
NEO-TOKYO HUB v4.1

STATUS:
LIVE STREAMING

SYSTEM HEALTH:
OPTIMAL

DATA FEED:
REAL-TIME



Kontakt(e)

Philipp Bellant

+43 660 651 44 01

philipp@aisberg.at

Rüdiger Linhart

+43 669 120 035 22

ruediger@aisberg.at

<https://www.aisberg.at>