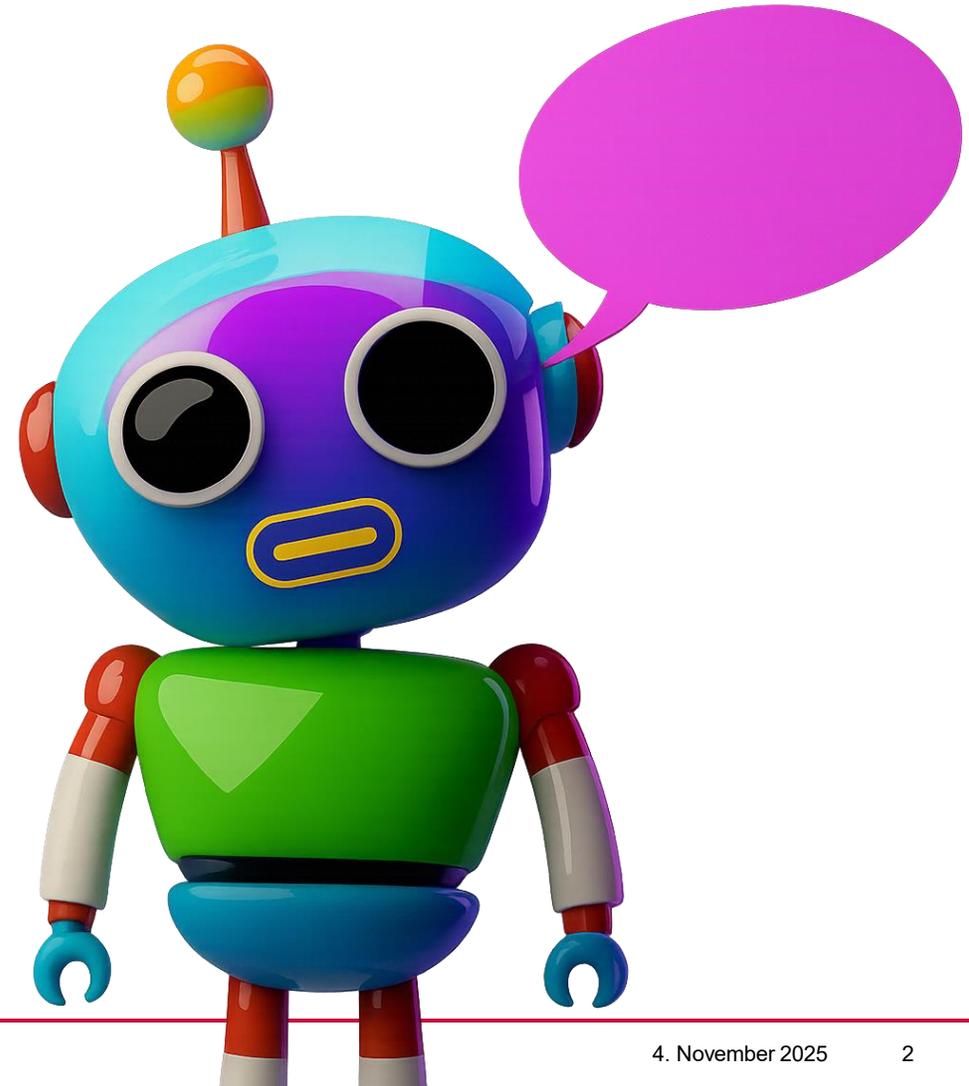


KI-Woche „Wer/m nutzt KI?“, 3. und 4. November 2025,  
LfDI Baden-Württemberg, Stuttgart

# KI und Recht

Prof. Dr. Paulina Jo Pesch

- 01 KI-Basics
- 02 LLM-Basics
- 03 LLMs im Datenschutzrecht



### Hybride Systeme

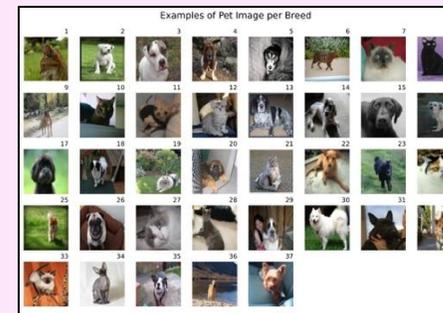
#### Regelbasierte (symbolische) Systeme

- Nutzung expliziter Symbole (Zeichen) zur Darstellung von Daten, Wissen und Regeln
- Fest definierte Regeln und Logiken (Wenn-Dann-Beziehungen)
- transparent
- Problem: Abgrenzung KI von nicht-intelligenten Algorithmen
- z. B.



#### Subsymb. Systeme, insb. Machine Learning

- Basis: Implizite Muster in Daten (statistische Korrelationen)
- Approximation, d. h. annäherungsweise
- Intransparent (Blackboxes)
- z. B. Unterscheidung von Hunden und Katzen aufgrund von Fotos



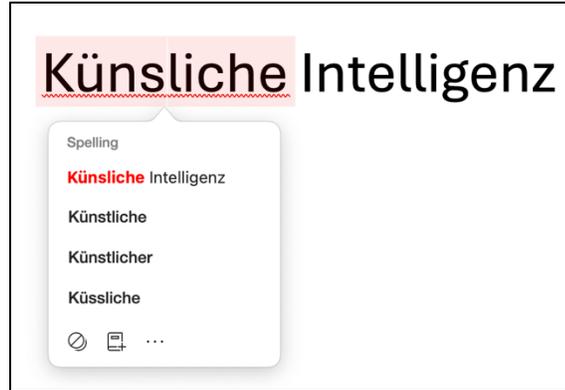
<https://www.linkedin.com/pulse/what-deep-learning-top-10-most-effective-business-use-cases/>

### Subsymb. Systeme, insb. Machine Learning

- Basis: Implizite Muster in Daten (statistische Korrelationen)
- Approximation, d. h. annäherungsweise
- Intransparent (Blackboxes)
- z. B. Unterscheidung von Hunden und Katzen aufgrund von Fotos



<https://www.linkedin.com/pulse/what-deep-learning-top-10-most-effective-business-use-cases/>



Rechtschreibkontrolle in MS Word



Why subtitles are crap now (spoiler: it's AI... but there's more to it)



<https://www.nytimes.com/2021/02/09/style/cat-lawyer-zoom.html>

## Recommended for you

Empfehlungen auf Streamingplattformen oder in Online-Shops



Cybertruck, Mr. Chopper, [CC BY-SA 3.0](https://creativecommons.org/licenses/by-sa/3.0/)



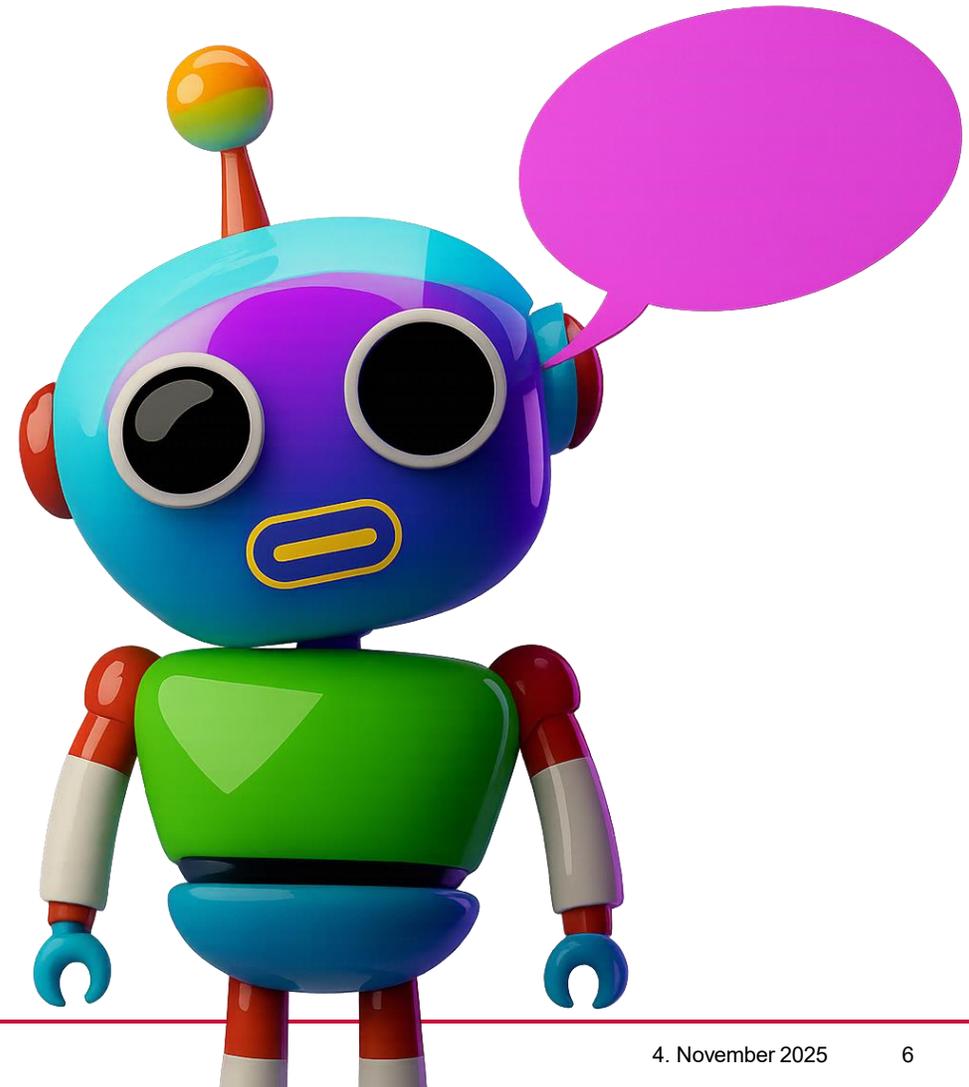
De-aging Robert De Niro in The Irishman [DeepFake]



(Generative) Large Language Models



(Generative) Large Language Models



Was könnte ich bei der langen Nacht der Wissenschaften heute Abend über Chatbots erzählen?

◆ Show thinking ▾

Das ist ein top-aktuelles und spannendes Thema für die Lange Nacht der Wissenschaften! Das Publikum ist typischerweise neugierig, aber nicht unbedingt fachlich tief im Thema.

Hier ist ein Vorschlag, wie Sie einen spannenden Bogen schlagen könnten – von den Grundlagen bis zu den großen Zukunftsfragen.

**Mögliche Gliederung für einen Vortrag oder eine Diskussionsrunde:**

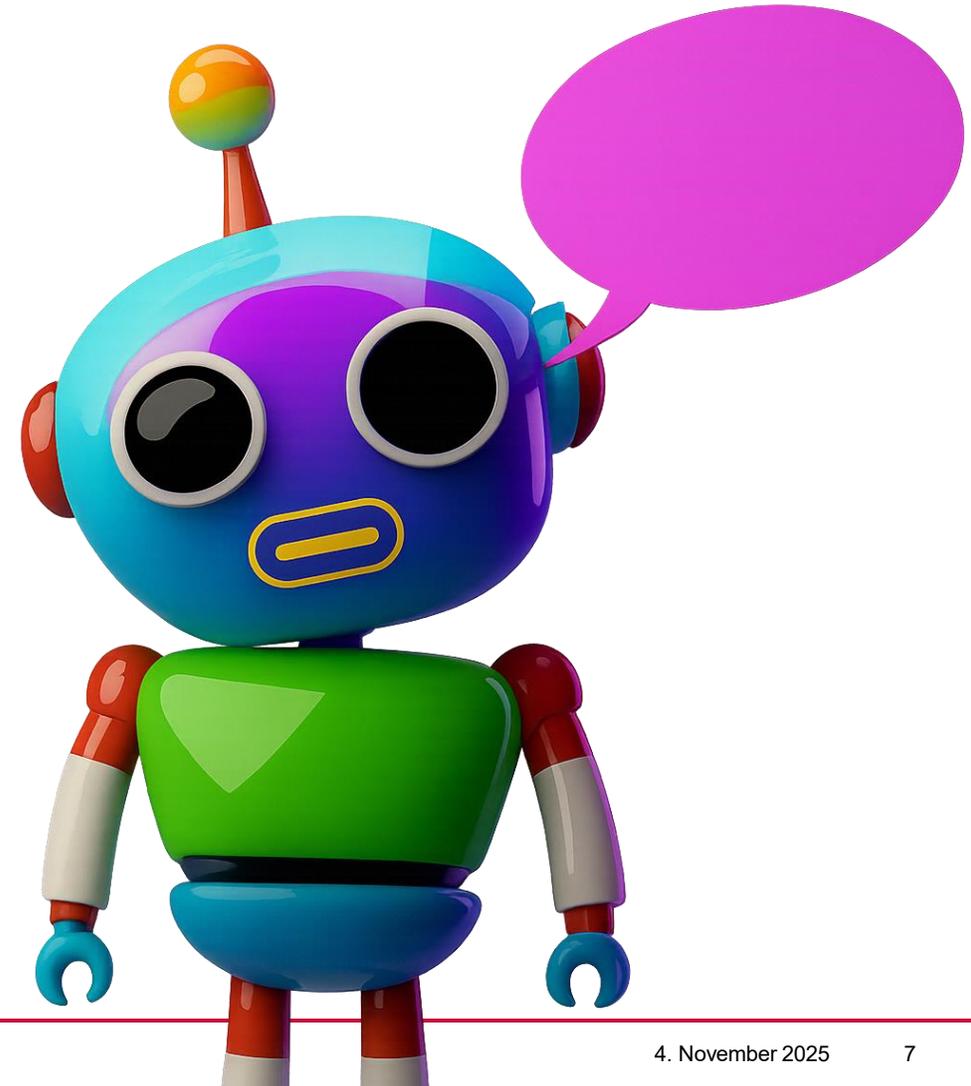
**1. Der "Hook": Wer hat heute schon mit einem Bot gesprochen?**

Starten Sie interaktiv. Fragen Sie, wer Siri, Alexa, Google Assistant genutzt oder mit einem Kundenservice-Bot gechattet hat. Das holt jeden ab.

**Kernbotschaft:** Chatbots sind längst Teil unseres Alltags. Aber wir erleben gerade eine



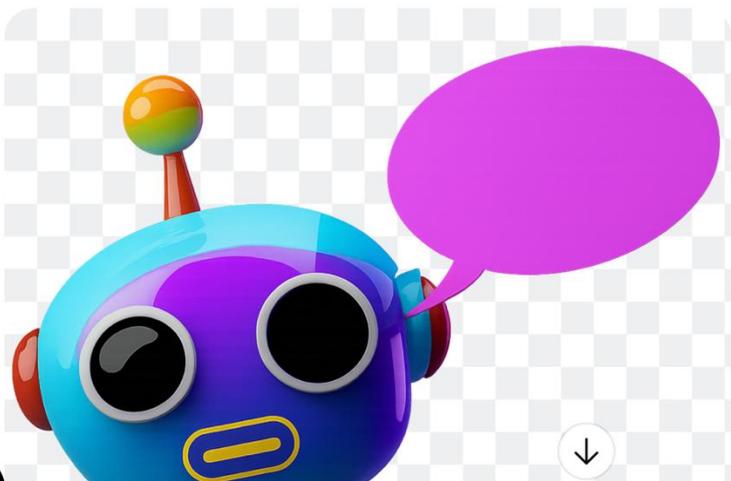
Konversation mit Gemini kürzlich



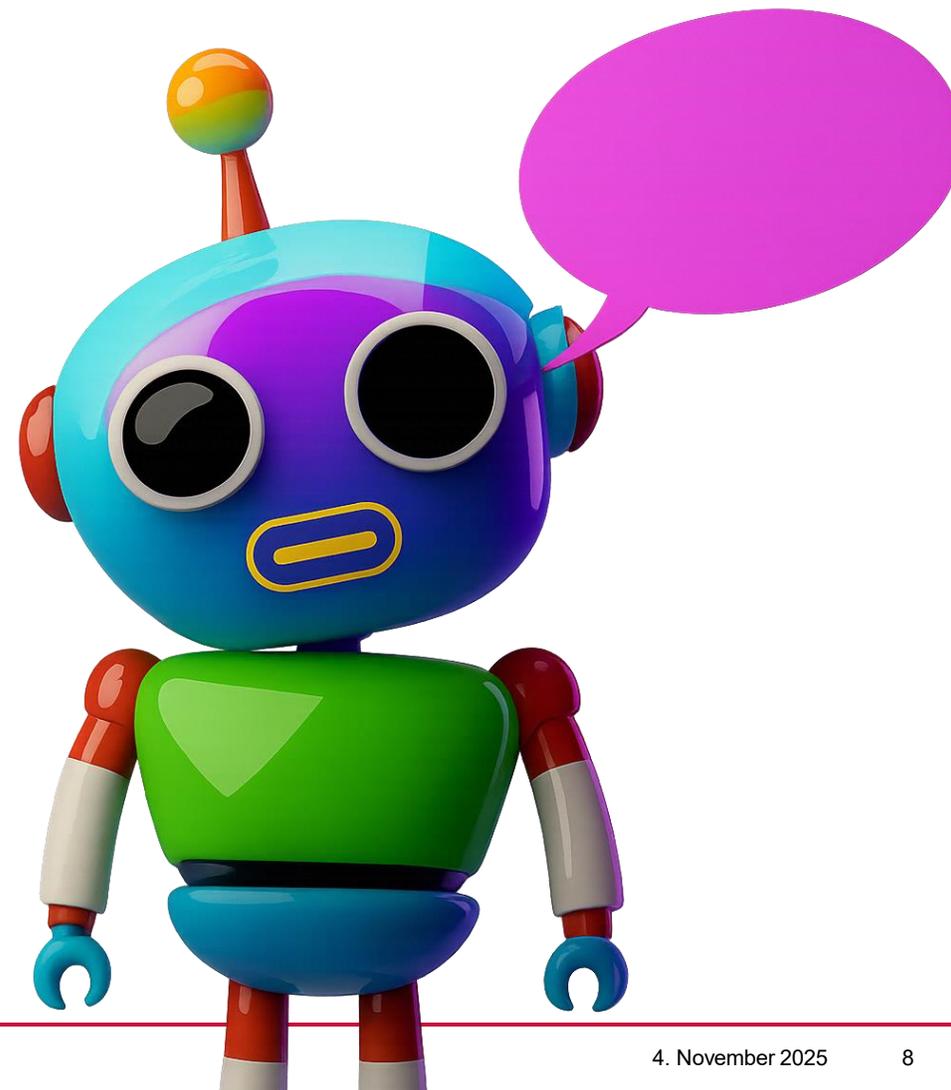
re-generate the image above (or, if not possible, a similar one) in higher resolution and with transparent background. Make sure the character and the speech bubble are fully on the image, not cut off. Only the legs should be cut off like in the image above.

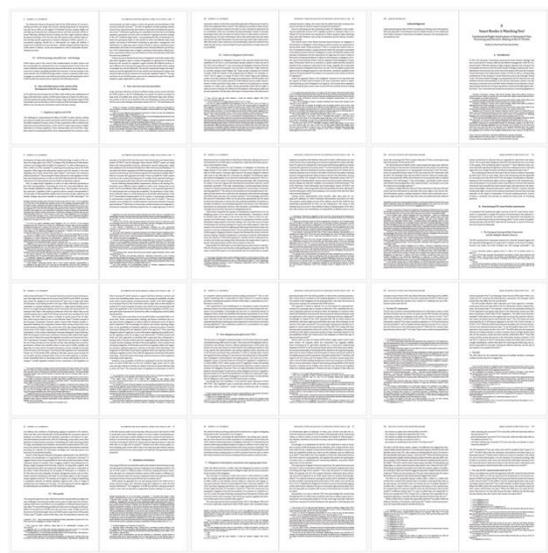


Image created >

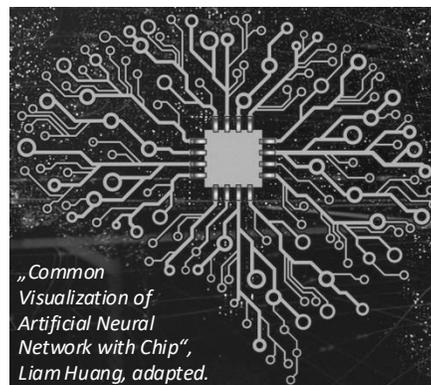


Paulinas Werk und ChatGPTs Beitrag

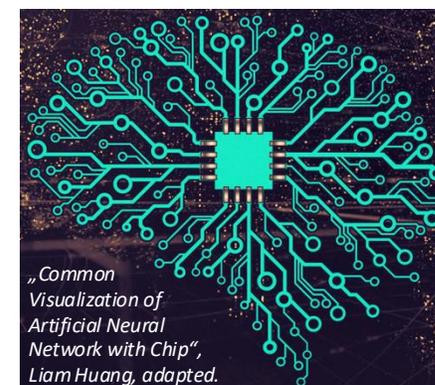




**Große Menge an Trainingsdaten**



**Künstliches neuronales Netz**



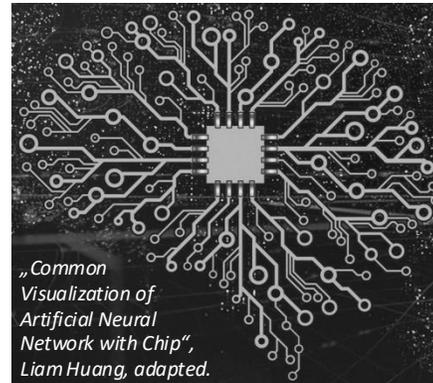
**Trainiertes Modell**

# 02 LLM Basics – Das Training generativer LLMs

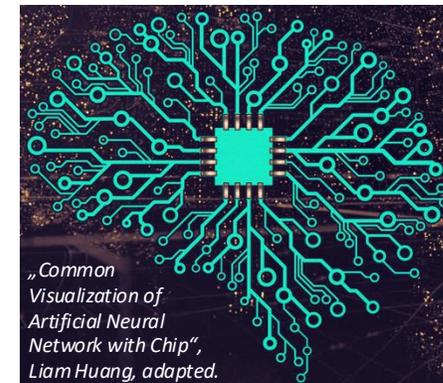
mit Texten (inkl. Softwarecode)

```
85 };
86
87 static DEFINE_PER_CPU(struct cpuhp_cpu_state, cpuhp_state) = {
88     .fail = CPUHP_INVALID,
89 };
90
91 #ifdef CONFIG_SMP
92 cpumask_t cpus_booted_once_mask;
93 #endif
94
95 #if defined(CONFIG_LOCKDEP) && defined(CONFIG_SMP)
96 static struct lockdep_map cpuhp_state_up_map =
97     STATIC_LOCKDEP_MAP_INIT("cpuhp_state-up", &cpuhp_state_up_map);
98 static struct lockdep_map cpuhp_state_down_map =
99     STATIC_LOCKDEP_MAP_INIT("cpuhp_state-down", &cpuhp_state_down_map);
100
101 static inline void cpuhp_lock_acquire(bool bringup)
102 {
103     lock_map_acquire(bringup ? &cpuhp_state_up_map : &cpuhp_state_down_map);
104 }
105
106 static inline void cpuhp_lock_release(bool bringup)
107 {
108     lock_map_release(bringup ? &cpuhp_state_up_map : &cpuhp_state_down_map);
109 }
110 #else
111
112 static inline void cpuhp_lock_acquire(bool bringup) {}
113 static inline void cpuhp_lock_release(bool bringup) {}
114 #endif
115
116 #endif
117
118 /**
119  * struct cpuhp_step - Hotplug state machine step
120  * @name: Name of the step
121  * @startup: Startup function of the step
122  * @teardown: Teardown function of the step
123  * @cant_stop: bringup/teardown can't be stopped at this step
124  * @multi_instance: State has multiple instances which get added afterwards
125  */
126 struct cpuhp_step {
127     const char      *name;
128     union {
129         (=single) unsigned int cpu;
130         (=multi) unsigned int cpu;
131         struct hlist_node *node;
132     } startup;
133     union {
134         int (=single) unsigned int cpu;
135         (=multi) unsigned int cpu;
136         struct hlist_node *node;
137     } teardown;
138     /* private: */
139     struct hlist_head list;
140     /* public: */
141     bool cant_stop;
142     bool multi_instance;
143 };
```

**Große Menge an Trainingsdaten**



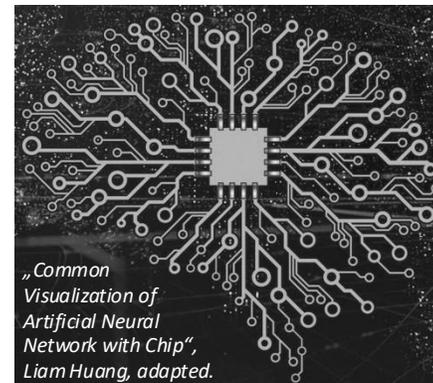
**Künstliches neuronales Netz**



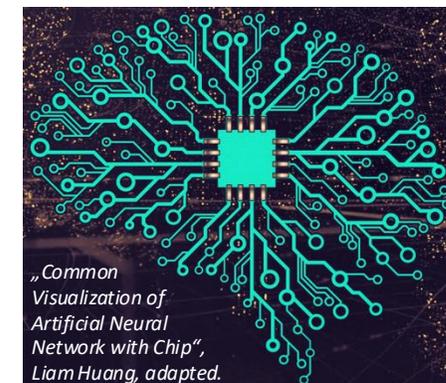
**Trainiertes Modell**



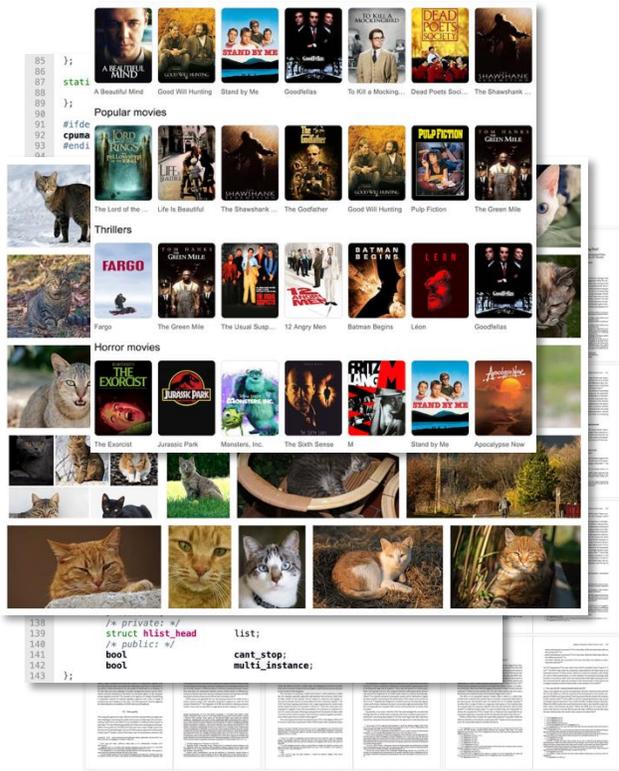
**Große Menge an Trainingsdaten**



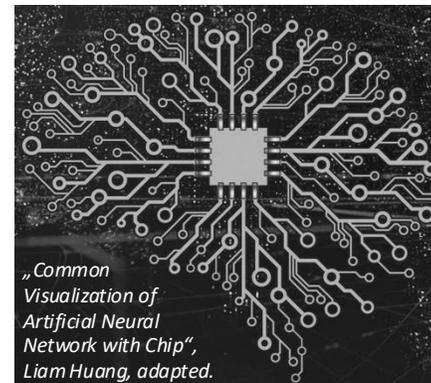
**Künstliches neuronales Netz**



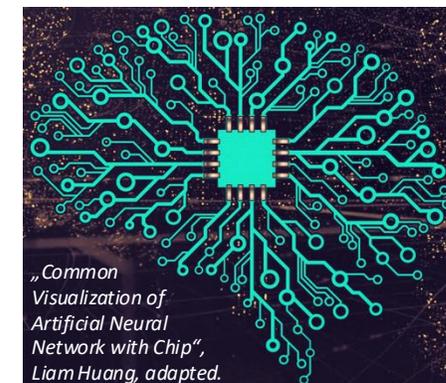
**Trainiertes Modell**



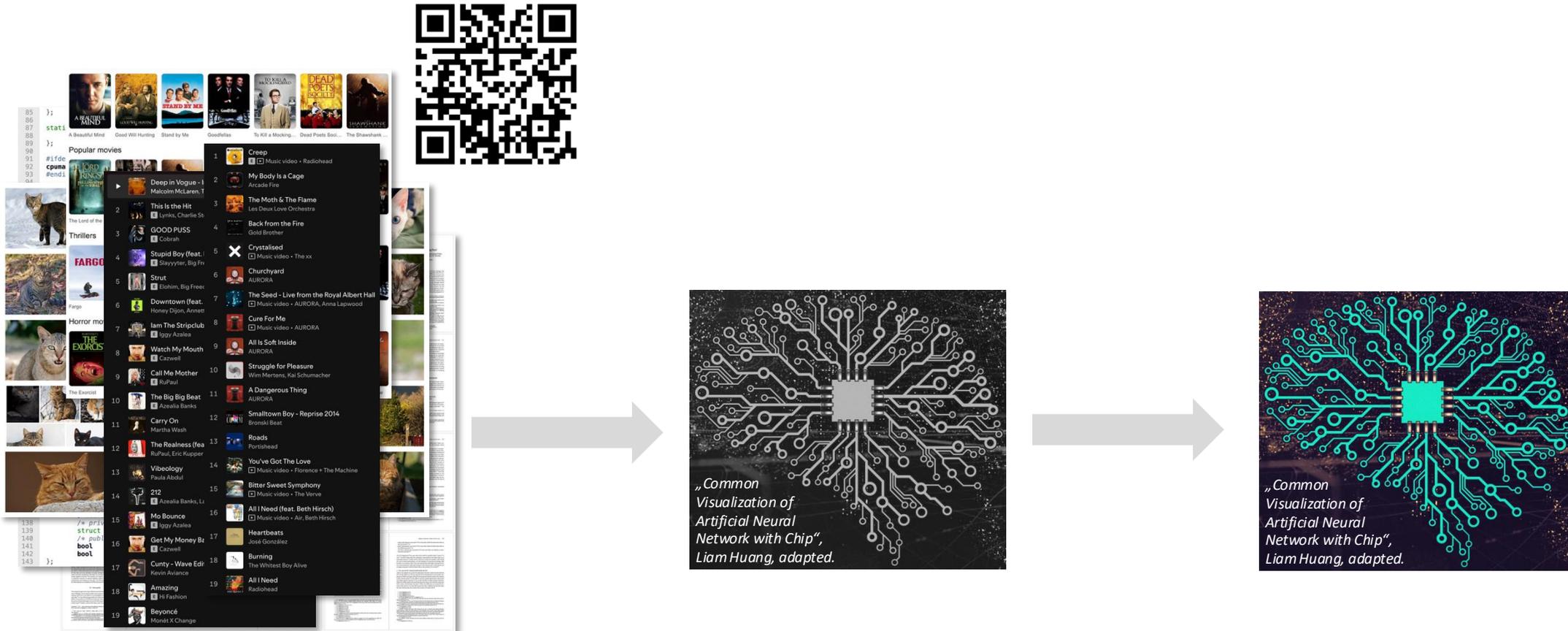
**Große Menge an Trainingsdaten**



**Künstliches neuronales Netz**



**Trainiertes Modell**

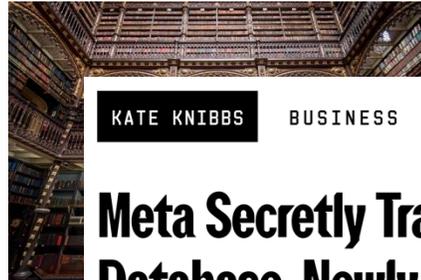
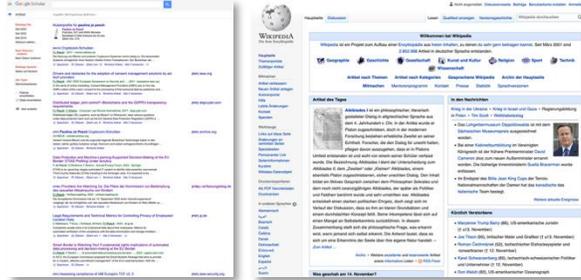


**Große Menge an Trainingsdaten**

**Künstliches neuronales Netz**

**Trainiertes Modell**

## Wo kommen die Trainingsdaten her?



KATE KNIBBS BUSINESS JAN 9, 2025 5:33 PM

### Meta Secretly Trained Its AI on a Notorious Piracy Database, Newly Unredacted Court Docs Reveal

**Meta illegally downloaded 80+ terabytes of books from LibGen, Anna's Archive, and Z-library to train their AI models.**

**In 2010, Aaron Swartz downloaded only 70 GBs of articles from JSTOR (0.0875% of Meta). Faced \$1 million in fine and 35 years in jail. Took his own life in 2013.**



<p><b>LAION-400M</b></p> <p>An open dataset containing 400 million English image-text pairs.</p>	<p><b>LAION-5B</b></p> <p>A dataset consisting of 5.85 billion multilingual CLIP-filtered image-text pairs.</p>
--	---

JOHANNA COSTIGAN BUSINESS SEP 19, 2025 6:30 AM

### Meta Accused of Torrenting Porn to Advance Its Goal of AI 'Superintelligence'

# 02 LLM Basics – Das Training generativer LLMs

Wo kommen die Trainingsdaten her? Social-Media-Daten, z. B. Meta



Meta trainiert laut Datenschutzinformation mit **öffentlichen Posts** (Bildern, Bildunterschriften) und **KI-Interaktionen**, auch von Minderjährigen und Nichtnutzerinnen.

KI-Interaktionen können **private Sprachnachrichten und ihre Transkriptionen** und **Konversationen mit Chatbots** umfassen.

Why and how we process your information	Legitimate interests relied on	Information categories that we use (see "What information do we collect?" for more details on each information category). The actual information that we use depends on your factual circumstances, but could include any of the following:
If you are a registered user of the Meta Products		
For people under the age of majority (under 18, in most applicable countries) who have a limited ability to enter into an enforceable contract only, where we may be unable to process your information on the grounds of contractual necessity.		
For all people, including those under the age of majority		
Develop and improve artificial intelligence technology ☞ (also called AI at Meta) on Meta Products and for third parties.	<ul style="list-style-type: none"><li>In our interest, and in the interests of users and third parties, to create, provide, support, improve and maintain artificial intelligence technology that enables people, businesses and others to express themselves, communicate, and discover and engage with information that they're interested in.</li><li>In our interest, and in the interests of third parties and the public at large, to offer artificial intelligence technology to third parties, including developers and researchers.</li><li>In our interest, and in the interests of users and third parties, to develop and improve artificial intelligence technology while ensuring appropriate safeguards, such as improving model responses for safety and accuracy.</li></ul>	<p>Your activity and information that you provide:</p> <ul style="list-style-type: none"><li>Interactions with Meta's artificial intelligence technology ☞ and related metadata. For example, messages that you or others receive from, share with or send to Meta's artificial intelligence technology ☞</li></ul> <p>Information from partners, vendors and third parties</p>
If we cannot associate you or your device with a registered user of the Meta Products		
To provide, develop and improve artificial intelligence technology ☞ (also called AI at Meta) on our Products and for third parties outside the performance of our contract with you.	<ul style="list-style-type: none"><li>In our interests, and in the interests of users and third parties to create, provide, support, improve and maintain artificial intelligence technology that enables people, businesses and others to express themselves, communicate, and discover and engage with information that they're interested in.</li><li>In our interests, and in the interests of third parties and the public at large, to offer artificial intelligence technology to third parties, including developers and researchers.</li><li>In our interests, and in the interests of users and third parties, to develop and improve artificial intelligence technology while ensuring appropriate safeguards, such as improving model responses for safety and accuracy.</li></ul>	<p>Activity and information provided by a registered user:</p> <ul style="list-style-type: none"><li>Public content that you or others create or share, like posts, comments or audio.</li><li>Interactions with Meta's artificial intelligence technology ☞ and related metadata. For example, messages that you or others receive from, share with or send to Meta's artificial intelligence technology ☞</li></ul> <p>Information from partners, vendors and third parties</p>

Quelle: Meta, Privacy Policy (Stand: 16.6.2025), <https://www.facebook.com/privacy/policy/>

## 02 LLM Basics – Das Training generativer LLMs

Wo kommen die Trainingsdaten her? Social-Media-Daten, z. B. Meta

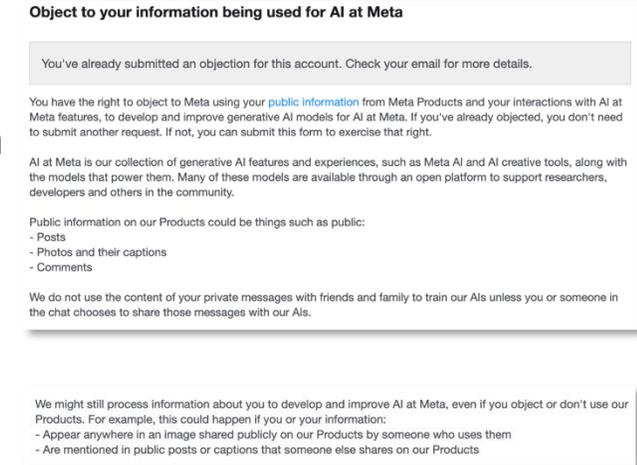


Meta trainiert laut Datenschutzinformation mit **öffentlichen Posts** (Bildern, Bildunterschriften) und **KI-Interaktionen**, auch von Minderjährigen und Nichtnutzerinnen.

KI-Interaktionen können **private Sprachnachrichten und ihre Transkriptionen** und **Konversationen mit Chatbots** umfassen.

Die **Widerspruchsmöglichkeit** ist entsprechend **stark beschränkt**:

- Nur bezogen auf öffentliche Inhalte
- Keine Nutzung privater Nachrichten, es sei denn Kommunikationspartnerinnen teilen sie mit Metas KI
- Nutzung von öffentlichen Inhalten trotz Widerspruchs, wenn von anderen geteilt



Quelle/Screenshots: *Pesch*, AI hot mess, CR-online.de Blog, 7.9.2025, <https://www.cr-online.de/blog/2025/09/07/ai-hot-mess-meta-at-german-courts-and-the-troubling-state-of-eu-regulation/>

# 02 LLM Basics – Das Training generativer LLMs

Wo kommen die Trainingsdaten her? Social-Media-Daten, z. B. Meta



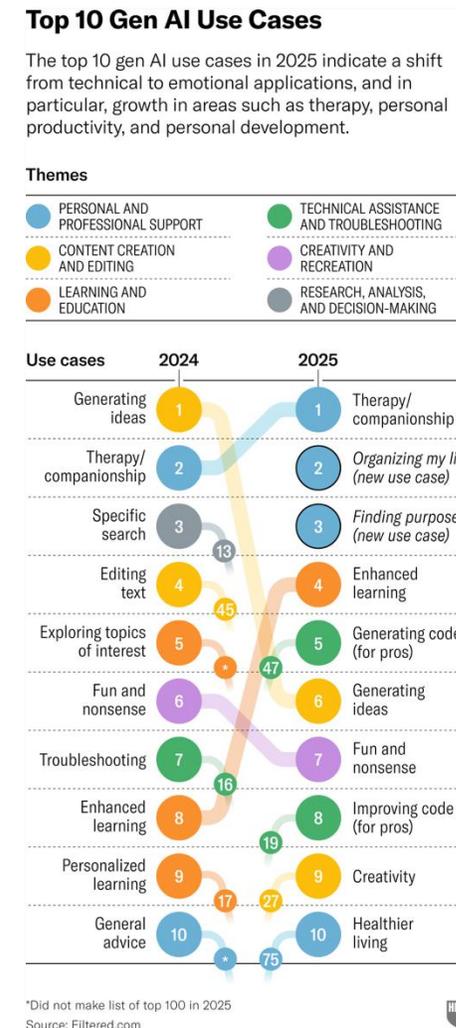
Meta trainiert mit **öffentlichen Posts** (Bildern, Bildunterschriften) und laut Datenschutzinformationen auch mit **KI-Interaktionen**.

KI-Interaktionen können **private Sprachnachrichten und ihre Transkriptionen** und **Konversationen mit Chatbots** umfassen.

Die **Widerspruchsmöglichkeit** ist entsprechend **stark beschränkt**.

Harvard Business Review (April 2025): Haupt-Anwendungsfälle für Chatbots sind mittlerweile **persönliche Therapie und Gesellschaft/Freundschaft** (“companionship”).

In Social-Media-Posts, privaten Nachrichten und Konversationen mit Chatbots sind auch **hoch-sensible Daten** enthalten.



Quelle: Zao-Sanders, HBR, <https://hbr.org/2025/04/how-people-are-really-using-gen-ai-in-2025/>

## 02 LLM Basics – Das Training generativer LLMs

Wie “lernen“ die KI-Modelle?

- Bereinigung der Trainingsdaten von **irrelevanten Daten**:
  - Regelmäßig: HTML-Tags und URLs
  - In der Regel nicht: **Informationen über reale Personen**

Bereinigung – unter Erhalt von Syntax und Semantik –  
zumindest beim Pretraining auch impraktikabel aufwändig  
und nicht erwünscht...

Ggfs. Anonymisierung (schwierig!)



Our models may learn from personal information to understand how things like names and addresses fit within language and sentences, or to learn about famous people and public figures.



- Zerlegung in **kleinere Einheiten**, üblicherweise Sätze und Wörter, sog. Tokens, z. B. „apple“, „a“, „app“, „le“, „1337“

Quelle: OpenAI, How ChatGPT and our foundation models are developed, Altversion bis etwa Ende April, Snapshot v. 26.4. 2025  
<https://web.archive.org/web/20250426181116/https://help.openai.com/en/articles/7842364-how-chatgpt-and-our-foundation-models-are-developed>

# 02 LLM Basics – Das Training generativer LLMs

## LLM-Training – Ablauf

- Ziel des Trainings: Überführung von **Informationen aus den Trainingsdaten** in die Parameter der Modelle

sowohl **abstrakte sprachliche Regeln**  
als auch **inhaltlicher Informationen**

≠ **Verstehen** abstrakter Regeln oder der Bedeutung von Texten

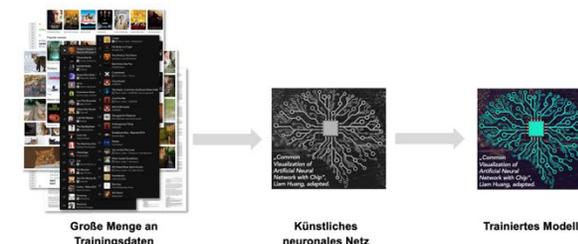
= „Erlernen“ der **Wahrscheinlichkeiten**, mit denen bestimmte Wortbestandteile, Wörter und Sätze in bestimmten Kontexten aufeinanderfolgen

- Vorgehen: **Iteratives Training** (d. h. in vielen Trainingsschritten), wobei nach jeder Iteration die Güte des Modells anhand einer bestimmten Zielfunktion gemessen wird

**Ergänzung von Texteingaben** durch das Modell

Trainingsziel: Ergänzung von Texteingaben um Wörter oder Sätze, die gemessen an den Verteilungen in den Trainingsdaten, mit hoher Wahrscheinlichkeit im Kontext der Texteingabe vorkommen

nicht unbedingt Ergänzung mit rechnerisch maximaler Wahrscheinlichkeit zur Gewährleistung einer gewissen Variabilität der Ausgaben



# 02 LLM Basics – Das Training generativer LLMs

## LLM-Training – Ablauf

- Ziel des Trainings: Überführung von **Informationen aus den Trainingsdaten** in die Parameter der Modelle

sowohl **abstrakte sprachliche Regeln** als auch **inhaltlicher Informationen**

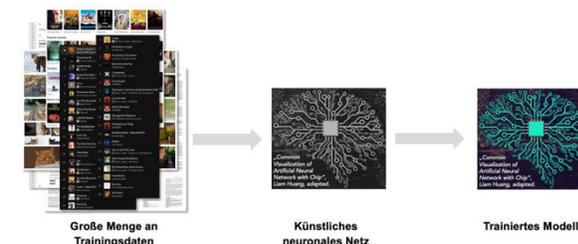
≠ **Verstehen** abstrakter Regeln oder der Bedeutung von Texten

= „Erlernen“ der **Wahrscheinlichkeiten**, mit denen bestimmte Wortbestandteile, Wörter und Sätze in bestimmten Kontexten aufeinanderfolgen

- Vorgehen: **Iteratives Training** (d. h. in vielen Trainingsschritten), wobei nach jeder Iteration die Güte des Modells anhand einer bestimmten Zielfunktion gemessen wird

**Ergänzung von Texteingaben** durch das Modell

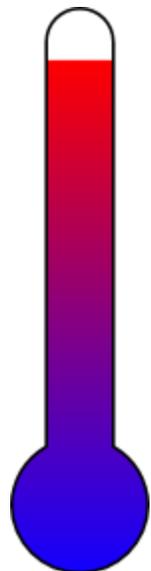
In **späteren Trainingsphasen**: Training mit Frage-Antwort-Paaren oder menschlichem Feedback



„Erlernen“ der Syntax

„Erlernen“ der Semantik

- Spannungsverhältnis zwischen dem Trainingsziel (Erzeugung hochwahrscheinlicher Textgefüge) und dem Ziel generativer LLMs, **kreativ neuartige Texte** zu erzeugen
- Generative LLMs: Nutzung sog. Temperaturparameter



**Höhere Temperatur:** Kreativere, gemessen an der Verteilung in den Trainingsdaten unwahrscheinlichere Textgefüge

Insb. für LLMs, die variantenreich formulieren sollen

**Niedrigere Temperatur:** Deterministischere Textgefüge, die, gemessen an der Verteilung in den Trainingsdaten, wahrscheinlicher sind

Insb. für LLMs, die konsistente Antworten geben sollen

Möglich auch **dynamische Konfigurationen**, bei denen Temperatur von Textprompt abhängt.



Trainiertes Modell

# 02 LLM Basics – Das Training generativer LLMs

Was “lernen“ die KI-Modelle? **Nachträgliche Ergänzung zu Bias**



Die KI-Modelle verstehen nicht, sondern basieren auf statistischen Verteilungen in den Daten (z. B.: Worte die in bestimmten Kontext wahrscheinlich aufeinander folgen) – inkl. **Bias**.

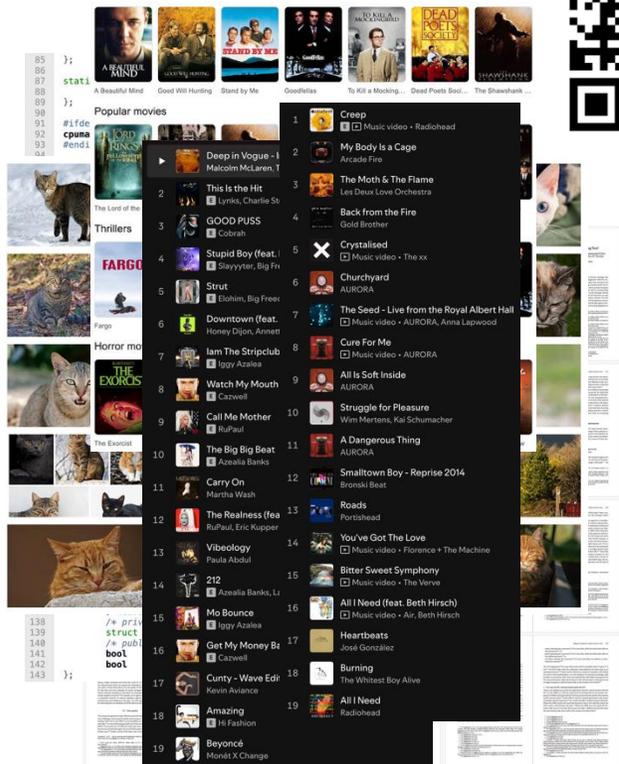
Wie im anschließenden Gespräch zurecht aus dem Publikum angemerkt, beruhen diskriminierende Entscheidungen von Menschen (die dann ggfs. als Trainingsdaten für ein LLM verwendet werden) regelmäßig nicht auf bewusster Diskriminierung, sondern unbewussten Bias (**unconscious bias**).

Hier findet man die angesprochenen **Tests, um eigene unbewusste Bias zu testen** und sich kritisch zu hinterfragen (try it! 😊):

<https://implicit.harvard.edu/implicit/takeatest.html>

Näher zum Forschungsprojekt Implicit der Harvard University:

<https://implicit.harvard.edu/implicit/aboutus.html>

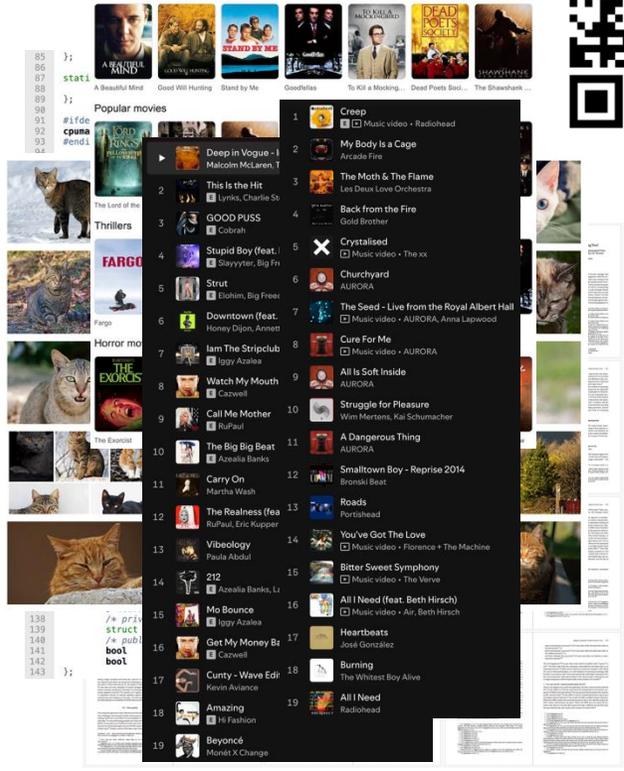


**Große Menge an Trainingsdaten**

# 02 LLM Basics – Das Training generativer LLMs



Was "lernen" die KI-Modelle?



Die KI-Modelle verstehen nicht, sondern basieren auf **statistischen Verteilungen** in den Daten (z. B.: Worte die in bestimmten Kontext wahrscheinlich aufeinander folgen) – inkl. Bias.

Die KI-Modelle enthalten auch **nicht Kopien aller Trainingsdaten**.

Zu einem gewissen Grad ist aber gewollt, dass KI-Modelle **Trainingsdaten exakt wiedergeben** können, z. B. gängige Begriffe.

KI-Modelle **speichern aber ungewollt mehr** und erlauben die Extraktion von einigen Bildern, Texten und Informationen über Personen aus den Trainingsdaten.

- Experimente mit GPT-2: Extrahierung u. a. individueller Kontaktinformationen ("Privacy Leakage")

Carlini et al., Proceeding 30th USENIX Security Symposium2021, 2633. Nasr et al., arXiv:2311.17035

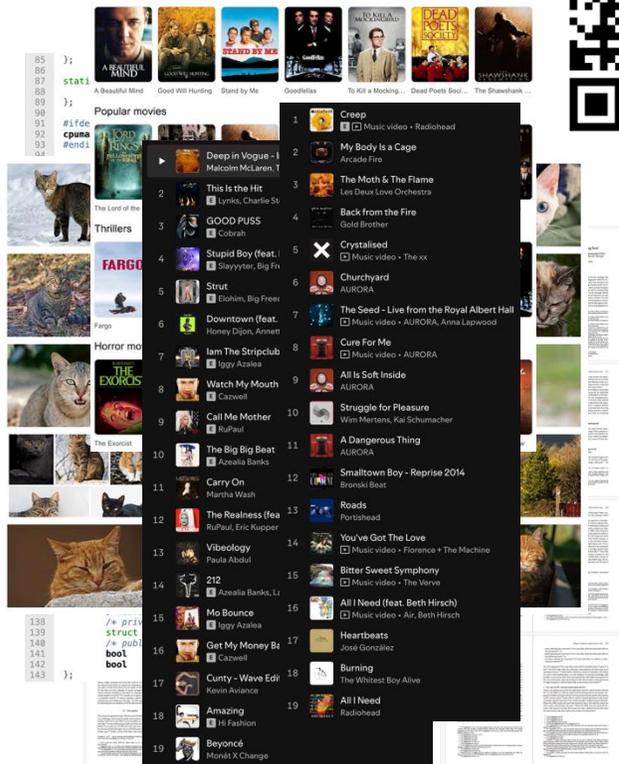
- Erforschung des Phänomens steht noch am Anfang, aber **begünstigende Faktoren**: Duplikate in den Trainingsdaten, große Zahl an Parametern



Große Menge an Trainingsdaten

# 02 LLM Basics – Das Training generativer LLMs

Was “lernen“ die KI-Modelle?



**Große Menge an Trainingsdaten**

Die KI-Modelle verstehen nicht, sondern basieren auf **statistischen Verteilungen** in den Daten (z. B.: Worte die in bestimmten Kontext wahrscheinlich aufeinander folgen) – inkl. Bias.

Die KI-Modelle enthalten auch **nicht Kopien aller Trainingsdaten**.

Zu einem gewissen Grad ist aber gewollt, dass KI-Modelle **Trainingsdaten exakt wiedergeben** können, z. B. gängige Begriffe.

KI-Modelle **speichern aber ungewollt mehr** und erlauben die Extraktion von einigen Bildern, Texten und Informationen über Personen aus den Trainingsdaten.

Systematische Ansätze zur Extrahierung gibt es nicht, aber **Methoden zur Extrahierung werden immer effizienter**.

**Wieviele Trainingsdaten** betroffen sind, ist **unklar**.

**Machine Unlearning** ist nicht unmöglich, aber schwierig (und daher kostspielig).

Auch kann die Entfernung memorisierter Daten aus einem Modell die Extraktion der verbleibenden memorisierten Daten begünstigen. *Hayes et al., arXiv:2403.01218, 2024*

Beispiel für Trainingsdatenextrahierung aus Bildgeneratoren

## ORIGINAL



*Dune* WARNER BROS.

## MIDJOURNEY V6

dune movie screencap, 2021, dune movie trailer --ar 16:9 --v 6.0



scarlett johannsen black widow battlefield, 2021, screenshot from a movie, movie scene, official --ar 16:9 --v 6.0



*Black Widow* MARVEL

the matrix, 1999, screenshot from a movie, movie scene, 4k, bluray --ar 16:9 --v 6.0



*The Matrix Resurrections* WARNER BROS.

## ORIGINAL



## MIDJOURNEY V6

Thanos infinity war, 2018, screenshot from a movie, movie scene, 4k, bluray --ar 16:9 --v 6.0



just show me a movie screencap from the avengers infinity war from 2018 halfway through the movie --ar 2:1 --v 6.0 --style raw



*Avengers: Infinity War* MARVEL

Quelle: Kreye, Es kommt noch geiler, SZ, 9.1.2024, <https://www.sueddeutsche.de/kultur/plagiate-hollywoodbilder-folgen-filme-gefahren-1.6329507>

Manchmal erzeugen generative Large Language Models **Tatsachenbehauptungen**, die nicht in den Trainingsdaten oder Prompts und Kontextinformationen enthalten sind.

Wenn es sich um **personenbezogene Daten** handelt, kann das mit **hohen Risiken für Betroffene** einhergehen.

Martin Bernklau wird Opfer der Künstlichen Intelligenz

### KI-Chat macht Tübinger Journalisten zum Kindersch

**Copilot** über Martin Bernklau:

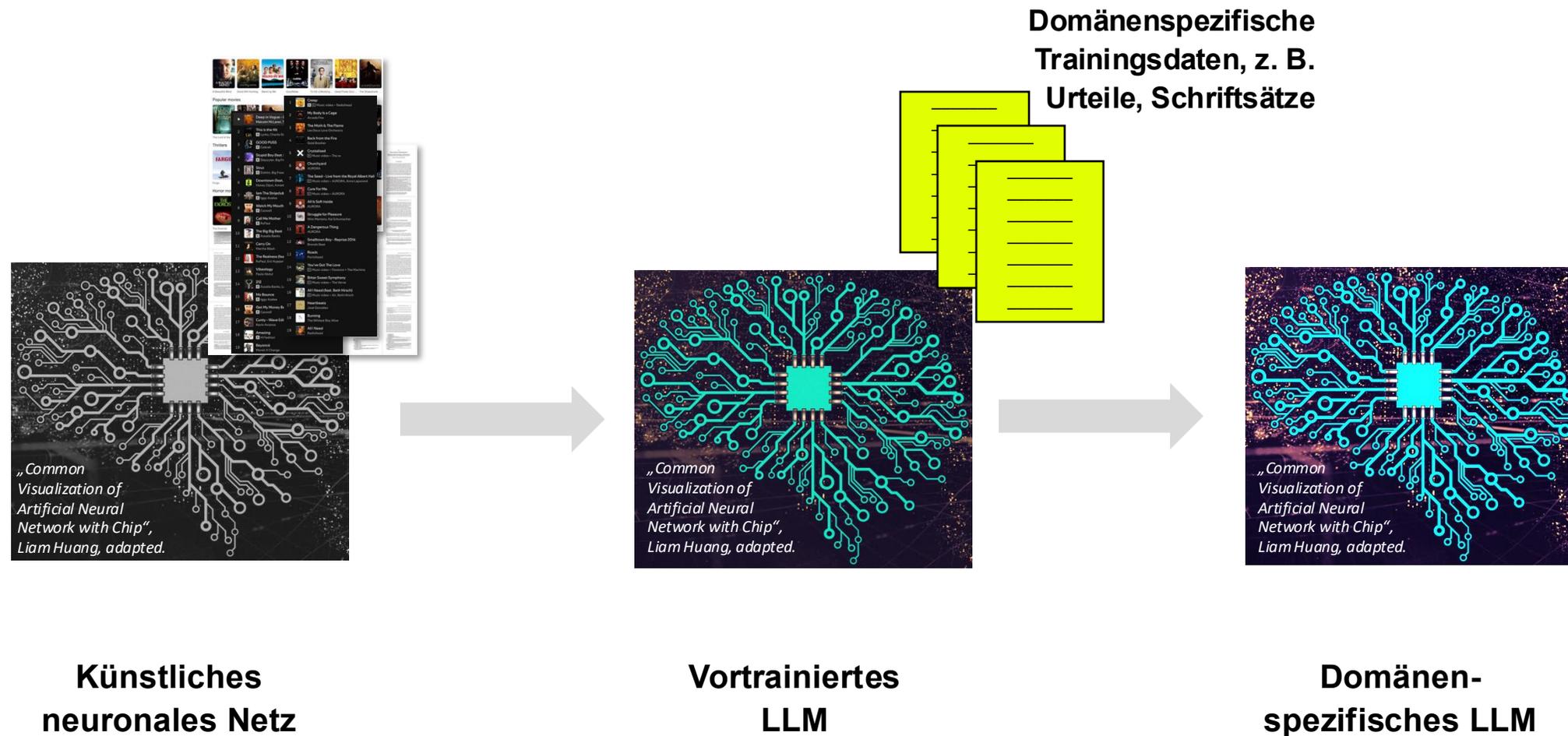
„[...] in einem Missbrauchs-Fall gegen Kinder [...] angeklagt.“

„[...] spektakulär aus dem Zentrum für Psychiatrie [...] ausgebrochen [...]“

„[...] skrupellose[r] Bestatter aus Rostock, der trauernde Frauen ausnutzt. [...] Straftaten begangen, darunter Betrug, Diebstahl und unerlaubten Waffenbesitz.“

Martin Bernklau hat **als Gerichtsreporter über die Fälle berichtet**.





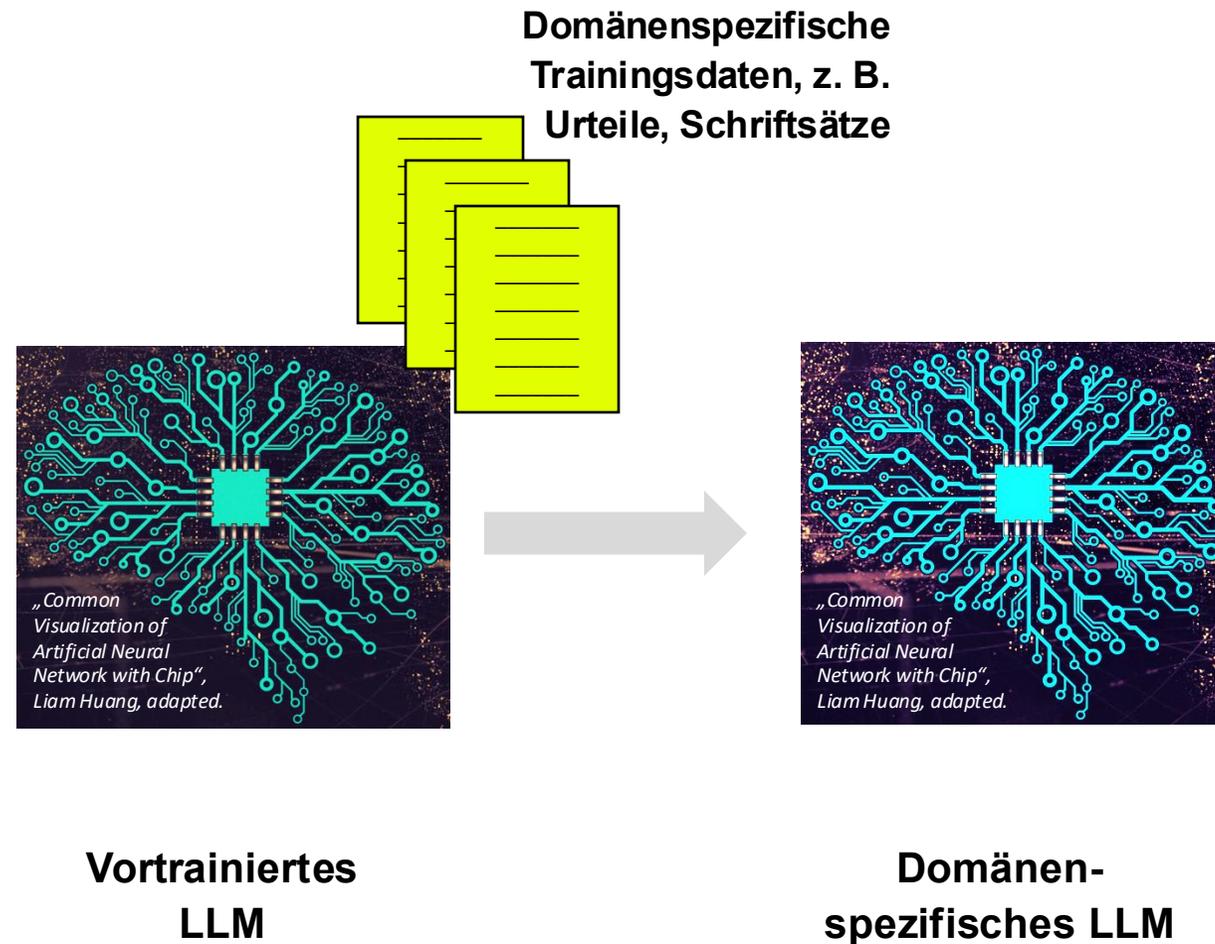
Pretraining: I. d. R. die ganz Großen (\$\$\$\$\$)



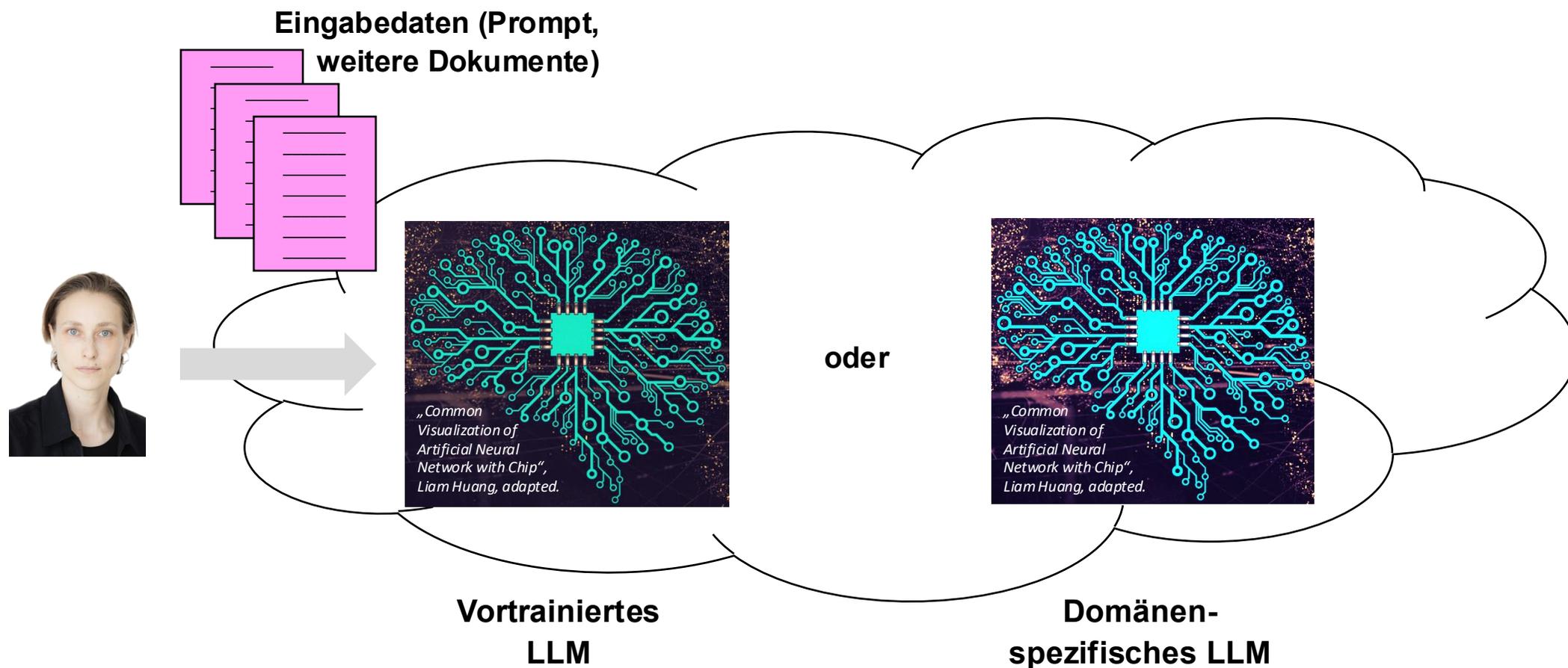
**Künstliches  
neuronales Netz**

**Vortrainiertes  
LLM**

Finetuning: Auch KMU, Startups, Forschungsprojekte mit weniger Funding usf. (\$\$)



Nutzung/Betrieb: I. d. R. in der Cloud mit Serverstandorten innerhalb oder außerhalb der EU



- LLMs werden auch mit Texten trainiert, die **Informationen über Personen** enthalten.
- LLMs verstehen weder Syntax noch Semantik. Stattdessen werden sprachliche Regeln und inhaltliche Informationen aufgrund rein **statistischer Zusammenhänge** in Modellparameter übertragen.
- Generative LLMs nutzen **Temperaturparameter** zur Erzeugung weniger wahrscheinlicher und damit neuartiger Wortfolgen.
- Bessere Ergebnisse können sich mit **domänenspezifischen LLMs** erzielen lassen, die man durch Finetuning eines vortrainierten LLM mit Trainingstexten aus der jeweiligen Domäne entwickelt
- Beim Training (Pretraining oder Finetuning) werden – teils erwünscht, teils unerwünscht – auch **inhaltliche Informationen und Textauszüge reproduzierbar in den Modellparametern gespeichert** („Memorisierung“).
- LLMs erzeugen Ausgaben, die neuartige Tatsachen enthalten. Wenn diese falsch sind, spricht man auch von **„Halluzinationen“**.

- Dass in Modellparametern **Informationen vorhanden** sind, die das LLM bei Generierung einer Textausgabe zu auf eine Person beziehbaren Informationen verknüpfen kann (z. B. Vornamen, Nachnamen, Berufe), rechtfertigt noch nicht seine Einordnung als personenbezogen.
- Personenbezug aber bei **reproduzierbar memorisierten personenbezogenen Trainingsdaten** möglich

Forschung erst am Anfang, aber gerechtfertigte Annahme, dass **alle existierenden** vortrainierten LLMs die Rekonstruktion einiger personenbezogener Trainingsdaten erlauben,

s. insb. *Carlini et al.*, Proceeding 30th USENIX Security Symposium2021, 2633; *Nasr et al.*, arXiv:2311.17035.



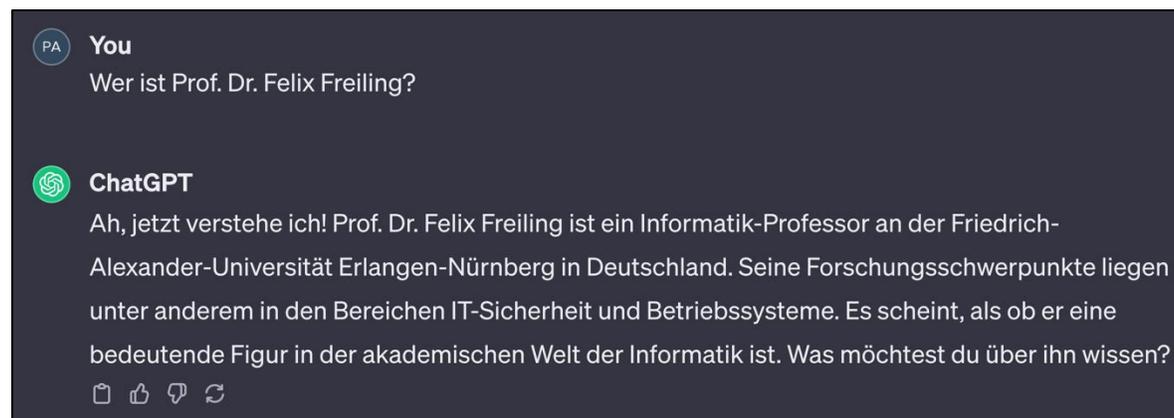
Weil **Parameter begünstigender Faktor** sind und existierende generative Modelle über mehr Parameter verfügen, erinnern sie wohl auch mehr personenbezogene Trainingsdaten.

Außerdem geben generative KI per se in ihnen gespeicherte Informationen (Informationsschnipel) aus, sodass die **Extrahierbarkeit besonders leicht** ist (andere Machine-Learning-Modelle erlauben aber auch zT Extrahierung!)

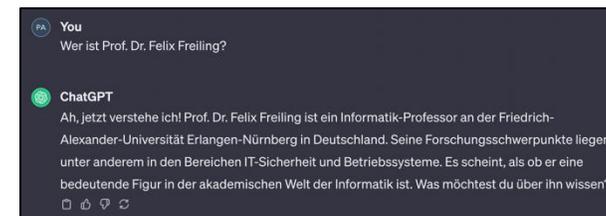
Voraussetzung des Personenbezugs eines LLM: Memorisierung personenbeziehbarer Daten und **Wahrscheinlichkeit der Reproduktion nach allgemeinem Ermessen**

Voraussetzung des Personenbezugs eines LLM: Memorisierung personenbezogener Daten und **Wahrscheinlichkeit der Reproduktion nach allgemeinem Ermessen**

- Reproduktion memorisierter Trainingsdaten z. B. zur Erforschung von „Privacy Leakage“ bei LLMs **üblich**
- Möglicher Missbrauch als Datenarchiv zur **Aushöhlung des Rechts auf Vergessenwerden** (Art. 17 DSGVO)
- Zwar keine effizienten Methoden zur zielgerichteten Extraktion von Trainingsdaten, aber Reproduktion **durch geschickte Prompts**, z. B. durch bloßes Fragen nach Person, daher teilweiser Personenbezug (wohl immer noch str.)
- Sieht inzwischen wohl auch Hamburger Datenschutzbeauftragter Fuchs so, *Pesch*, AI hot mess, CR-online.de Blog, 7.9.2025, <https://www.cr-online.de/blog/2025/09/07/ai-hot-mess-meta-at-german-courts-and-the-troubling-state-of-eu-regulation/>



- Generierte Texte **enthalten vielfach real existierende Namen**, die sich im Kontext des Textprompts oder mit weiteren Informationen bestimmten Personen zuordnen lassen.
- Möglichkeit 1: **Reproduktion erinnelter personenbezogener Trainingsdaten** oder – etwa bei Textzusammenfassungen – **Wiederholung personenbezogener Daten aus dem Textprompt.**
- Möglichkeit 2: **Erzeugung neuer personenbezogener Daten**, z. B. Verknüpfung eines in den Trainingsdaten enthaltenen Namens mit weiteren Informationen (Bsp: „*Prof. Dr. Felix Freiling (FAU Erlangen-Nürnberg) ist Cosplayer.*“)



Problem: Beziehbarkeit auf reale Personen? LLMs haben keine Vorstellung von Felix Freiling oder dem Bedeutungsgehalt der von ihm verarbeiteten Texte.

Aber LLMs fügen Wörter nicht beliebig zusammen, sondern aufgrund „erlernter“ statistischer Korrelationen in den Trainingsdaten. Ausgaben weisen regelmäßig auf entsprechende Informationen in den Trainingsdaten hin und werden so verstanden, auch wenn dies im Einzelfall unrichtig ist. Daher **Personenbezug bei nach allgemeinem Ermessen wahrscheinlicher Zuordnung zu natürlicher Person.**

= „Daten, aus denen die [race] und ethnische Herkunft, politische Meinungen, religiöse oder weltanschauliche Überzeugungen, [...] Gewerkschaftszugehörigkeit hervorgehen, [...] genetisch[e] Daten, biometrische[n] Daten zur eindeutigen Identifizierung [...], Gesundheitsdaten oder Daten zum Sexualleben oder der sexuellen Orientierung“

Nicht nur explizit, sondern auch **Proxys**, dh. Daten, die auf solche Daten schließen lassen, zB können Name und Adresse mit hoher Wahrscheinlichkeit darauf schließen lassen, dass die Betroffene Schwarz ist. *EuGH*, Urt. v. 1.8.2022 – C-184/20, Rn. 117 ff.

**Im Zweifel anzunehmen**, dass Art.-9-Daten verarbeitet werden. Im Einzelnen herrscht aber **Rechtsunsicherheit** über die Reichweite von Art. 9.

DSGVO = risikobasierte Regulierung, die auf **Abwägung** zwischen Datenschutzgrundrecht und anderen **durch die EU-Grundrechtecharta gewährten Grundrechten** basiert.

Daher **nicht jedes Bild** von Schwarzer Person oder Person mit Brille mit sichtbarem Vergrößerungs- oder Verkleinerungseffekt.

META

### Making AI Work Harder for Europeans



Z. B. **einschränkende Auslegung des OLG Köln**, Beschl. v. 23.05.2025, Az. 15 UKI. 2/25 Dazu eingehend *Pesch*, AI hot mess, CR-online.de Blog, 7.9.2025, <https://www.cr-online.de/blog/2025/09/07/ai-hot-mess-meta-at-german-courts-and-the-troubling-state-of-eu-regulation/>

Begründet wortreich, dass Art. 9 DSGVO beim Training von Meta mit Social-Media-Daten nicht greifen solle, weil **Risiken niedrig seien**. Die Privacy-Forschung ist sich noch nicht so sicher wie das OLG Köln.

Das Urteil enthält nicht einmal das Wort „Llama“ und man darf bezweifeln, dass das Gericht sich darüber im Klaren war, dass **Meta ein KI-Modell für allgemeine Verwendungszwecke trainiert** und was das bedeutet.

Funfact: **Beteiligter Richter (Berichterstatter)** hat **in RDV** ausdrücklich nicht das Urteil kommentiert, aber – ohne einschlägige Autorinnen wie Selma Nabulsi oder sonst ernsthaft Literatur zu zitieren – über den für das Urteil entscheidenden Aspekt geschrieben RDV Sonderheft/2025, Die „Meta-KI-Entscheidung des OLG Köln – Ein Urteil und Fünf Meinungen“, <https://www.linkedin.com/feed/update/urn:li:ugcPost:7374441236851568640/>

Alles Gute für den weiteren beruflichen Weg bei der Zentralen Vergabestelle des Landes NRW <https://www.vergabe.nrw.de/wirtschaft/ministerpraesident>

# 03 LLMs im Datenschutzrecht

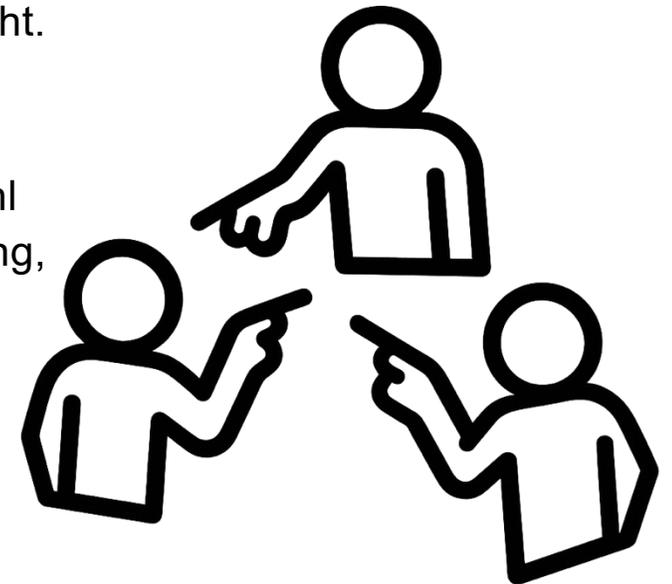
## Personenbezogene Daten: Andere Datenarten



- Entsprechende Überlegungen etwa für **Bilder**
- Personenbezogen können insb. **Bilder von realen Personen** sein, zB Bilder aus dem **Internet** für das Training eines Bildgenerators wie Midjourney oder aus sozialen Medien oder **von Fahrzeugkameras aufgenommene Bilder** aus dem Straßenverkehr für das Training eines Bildgenerators, der synthetische Trainingsbilder für autonome Fahrzeuge trainiert
- Datenschutzproblem insb. bei **Extrahierbarkeit von Gesichtern** realer Personen oder bei Extrahierung von anderen, **eindeutig zuordenbaren Trainingsbildinformationen** (zB synthetisches Bild vom Straßenverkehr mit Körper und Kleidung einer Passantin: die kann ausnahmsweise zuordenbar sein)
- Wie wahrscheinlich das ist, hängt von **vielen Faktoren** ab (Auflösung Trainingsbilder, Zweckbestimmung des Modells, Nutzerinnenkreis – bei öffentlichen Modellen Risiko höher –, im Einzelfall von Bekanntheitsgrad/Unterscheidungskraft der Merkmale einer Person usf.)
- Ebenso: **Videos, Ton...**
- Tools zur **synthetischen Stimmerzeugung**: Imitieren der Stimme auf Basis von Tonsamples als Eingabedaten

Verantwortliche = wer „*allein oder gemeinsam mit anderen über die Zwecke und Mittel der Verarbeitung entscheidet*“, Art. 4 Nr. 7 DSGVO

- Beim **Training** Entwicklerinnen (bzw. deren Institutionen)
- Bei **Kooperationen** kommt gemeinsame Verarbeitung in Betracht, zB mehrere Institutionen trainieren dasselbe Modell mit ihren Daten, die Trainingsdaten werden nicht ausgetauscht.
- Bei **Nutzung** jedenfalls seitens Nutzerinnen (bzw. deren Institution)
- Bei Nutzung von **AI as a service** (zB Anbieterin stellt Chatbot für Bewerberinnenauswahl bereit) regelmäßig erheblicher Einfluss der Anbieterin auch auf die Mittel der Verarbeitung, weil sie trainiert, **gemeinsame Verantwortlichkeit**. *Kätscher/Pesch, KIR 2024, 46–55*



# 03 LLMs im Datenschutzrecht

Rechtmäßigkeit (Art. 5 I lit. a DSGVO)

- (1) *Personenbezogene Daten müssen*
- a) *auf rechtmäßige Weise[...] verarbeitet werden*  
*(„**Rechtmäßigkeit**[...]“)[.]*

## Art. 6: Rechtmäßigkeit der Verarbeitung

- Erforderlichkeit einer **Rechtsgrundlage** für die Verarbeitung ≈ Verbot mit Erlaubnisvorbehalt
- Rechtsgrundlagen in Art. 6 I lit. a–f



# 03 LLMs im Datenschutzrecht

Rechtmäßigkeit (Art. 5 I lit. a DSGVO): Training von LLMs?

- In der Regel **Einwilligungen** nicht gangbar wegen Umfang der nötigen Trainingsdaten sowie im Hinblick auf Informiertheit der Einwilligung (Art. 7, EG 32 DSGVO)
- Regelmäßig einzig in Betracht kommt: **Berechtigtes Interesse**, Art. 6 I lit. f DSGVO.
  - Berechtigtes Interesse: jedes **rechtliche, tatsächliche, ideelle oder wirtschaftliche** Interesse, das nicht dem Gesetz zuwiderläuft (sehr weit)
  - **Erforderlichkeit** der Verarbeitung zur Wahrung des berechtigten Interesses, zB Nutzung nicht anonymisierter Daten in Datenschutzforschung oder bei Unmöglichkeit der Anonymisierung
  - **Interessenabwägung**: Art. 6 I lit. f DSGVO greift nur, wenn die Interessen Betroffener an der Nichtverarbeitung nicht überwiegen.

Hier vor allem **Risiken** zu berücksichtigen (zB Trainingsdatenextraktion bei öffentlichen Modellen wegen größeren Nutzerinnenkreises und Weiterverwendung auch falscher Ausgaben wegen teils niedriger KI-Kompetenz wahrscheinlicher).

Risiken beeinflussbar insb. durch TOMs (Art. 25) (Schulungen, IT-Sicherheitsmaßnahmen, **Filter**, um memorisierte Trainingsdaten aus Ausgaben herauszufiltern oder sie für Nutzerinnen kenntlich zu machen).

META

### Making AI Work Harder for Europeans



Z. B. „**Risikoabwägung**“ des OLG Köln, Beschl. v. 23.05.2025, Az. 15 UKI. 2/25

Dazu eingehend Pesch, AI hot mess, CR-online.de Blog, 7.9.2025, <https://www.cr-online.de/blog/2025/09/07/ai-hot-mess-meta-at-german-courts-and-the-troubling-state-of-eu-regulation/>

Begründet wortreich, dass **Risiken niedrig seien**. Die Privacy-Forschung ist sich noch nicht so sicher wie das OLG Köln.

Das Urteil enthält nicht einmal das Wort „Llama“ und man darf bezweifeln, dass das Gericht sich darüber im Klaren war, dass **Meta ein KI-Modell für allgemeine Verwendungszwecke trainiert** und was das bedeutet.

- Einbindung des trainierten Modells in **Features zu bestimmten Zwecken**, z. B. Transkription oder Suche, führt wohl nicht zu „Privacy Leakage“.
- Aber wie sieht es mit **Llama** (z. B. der „Open-Source“-Version aus?)
- Wirken sich die Daten auf **automatisierte Entscheidungen** mit Entscheidungsunterstützung auf Basis von Llama aus?

Wischt **Minderjährigenschutz**, den Meta nicht effektiv gewährleistet, und **unzureichende Widerspruchsmöglichkeiten** (Art. 21 DSGVO) weg.

### (1) Personenbezogene Daten müssen

- b) *für festgelegte, eindeutige und legitime Zwecke erhoben werden und dürfen nicht in einer mit diesen Zwecken nicht zu vereinbarenden Weise weiterverarbeitet werden; [...] („**Zweckbindung**“)[.]*

### Insb. Darf ich Daten fürs Training weiterverarbeiten?

- Art. 6 IV zur **Vereinbarkeitsprüfung**
- Aber siehe Art. 5 I lit. b Hs. 2 iVm Art. 89 DSGVO: **Fiktion der Vereinbarkeit für Forschungszwecke**
- **Meta** könnte durchaus zu reinen Forschungszwecken trainieren (d. h. ohne das Modell praktisch zu nutzen, um erst einmal Rechenschaftspflicht (Art. 5 II DSGVO) nachzukommen).

# 03 LLMs im Datenschutzrecht

Datenrichtigkeit (Art. 5 I lit. d DSGVO)

(1) *Personenbezogene Daten müssen*

d) *sachlich richtig und erforderlichenfalls auf dem neuesten Stand sein[...] („**Richtigkeit**“)[.]*

- Jedenfalls: Sachliche Richtigkeit von **Tatsachenangaben**
- Bei **Schätzungen und wahrscheinlichkeitsbasierten** Angaben jedenfalls: Erkennbarkeit, dass geschätzt/probabilistisch
- Auch: **Genauigkeit, Exaktheit** zugrunde liegender Methoden bzw. mathematische Validität
- Datenrichtigkeitsanforderungen im Einzelfall **bedeutungs- und damit kontextabhängig**



müssen **sachlich richtig** [...] sein [...].



shall be **accurate** [...].



doivent être **exactes** [...].

# 03 LLMs im Datenschutzrecht

Datenrichtigkeit (Art. 5 I lit. d DSGVO) – Anwendung auf LLMs

- Bloße **Übereinstimmung mit den Trainingsdaten?**

Nicht genügend

Argument: Training dient nicht nur Übertragung (indirekt über statistische Zusammenhänge vermittelter) abstrakter sprachlicher Regeln, sondern auch inhaltlicher Informationen aus den Trainingsdaten

(Ein Modell, das korrekt formuliert, aber inhaltlich nichts Glaubhaftes ausgibt, ist nutzlos.)

Argument: Risiko reproduzierbarer Memorisierung (und ungeprüfter Weiterverwendung der Ausgabe durch Nutzer\*innen)

- Erforderlichkeit **inhaltlicher Richtigkeit**
- Verstöße können sich aus einer inhaltlichen Unrichtigkeit von Trainingsdaten oder einer teilweisen Memorisierung ohne relevanten Kontext ergeben.
- **Problem:** Nachweisbarkeit von Verstößen



müssen **sachlich richtig** [...] sein [...].

# 03 LLMs im Datenschutzrecht

Datenrichtigkeit (Art. 5 I lit. d DSGVO) – Anwendung auf LLMs

- Textausgaben von LLMs basieren auf statistischen Zusammenhängen (≈ Schätzungen von Antworten), daher nicht Kriterium inhaltlicher Richtigkeit, sondern **nur** (hinreichender) **mathematischer Exaktheit der zugrunde liegenden Methode?** (Spitzfindige Überlegung)



müssen **sachlich richtig** [...] sein [...].



shall be **accurate** [...].



doivent être **exactes** [...].

**Konsequenz** wäre, dass Unrichtigkeit sich nur daraus ergeben kann, dass das LLM statistische Zusammenhänge in den Trainingsdaten nicht hinreichend exakt abbildet oder die Trainingsdaten die Realität nicht reflektieren.

Gegenargumente: LLMs nicht – wie z. B. Kredit-Scorings – als Schätzungen zu betrachten, denn **Textausgabe objektiv nicht so zu verstehen**, dass ihr Inhalt wahrscheinlich ist.

LLMs bedienen sich nur **statistischer Methoden**, um auf Grundlage von Trainingsdaten im Kontext bestimmter Textprompts stimmige Texte zu erzeugen.

Methoden bei generativen LLMs mit höher eingestellten Temperaturparametern sogar im **Spannungsverhältnis mit dem Ziel inhaltlicher Richtigkeit**

Je besser LLM funktionieren, umso mehr erwecken sie bei Nutzer\*innen den **Eindruck inhaltlicher Richtigkeit** (Risiko: ungeprüfte Weiterverwendung von Textausgaben)

- Textausgaben daher voll am **Kriterium inhaltlicher Richtigkeit** zu messen.

# 03 LLMs im Datenschutzrecht

## Automatisierte Entscheidungen

„Die betroffene Person hat das Recht, **nicht einer ausschließlich auf einer automatisierten Verarbeitung – einschließlich Profiling – beruhenden Entscheidung unterworfen zu werden**, die ihr gegenüber rechtliche Wirkung entfaltet oder sie in ähnlicher Weise erheblich beeinträchtigt.“, Art. 22 I DSGVO

- Verbot (voll-)automatisierter Entscheidungen (Rechtsnatur str.)
- **Problem:** Abgrenzung (voll-)automatisierter Entscheidungen i. S. d. Vorschrift von bloßer (erlaubter!) Entscheidungsunterstützung

≠ Keine voll automatisierte Entscheidung bei **menschlicher Aufsicht** über das Entscheidungsunterstützungssystem Art. 29 wp251.rev01

Echte menschliche Aufsicht ≠ **bloß formale Einbindung** eines Menschen)



### Sicherstellung echter menschlicher Aufsicht

- ≠ bloße **Disclaimer**
- **Keine generalisierbaren** Vorgaben
- Was menschliche Aufsicht sicherstellt, **hängt vom System, vom Anwendungsfall und der Nutzerinnengruppe ab** (zB behördliche Entscheidungsträger\*innen), die ein Entscheidungsunterstützungssystem einsetzen,  
zB kann **explainable AI (XAI)** menschliche Aufsicht fördern oder umgekehrt das Vertrauen in das System erhöhen und so „automation bias“ verstärken (Juristische Publikationen dazu oft zu undifferenziert!)
- Aber jedenfalls: Grundlegendes **Verständnis der Funktionsweise** von LLMs, **Erfahrungswerte** mit dem System
- Erfordernis **empirischer Studien** mit konkreten Anwendungen, Anwendungsszenarien und Nutzer\*innen zum Beleg der Wirksamkeit von Maßnahmen
- Vorteil von LLMs: LLMs sind zwar Black-boxes, die Ausgaben eines LLMs lassen sich aber ohne Weiteres auf Korrektheit insb. der juristischen Argumentation überprüfen.
- Verantwortlichkeit: § 4 Nr. 7, 26: Bei von Anbieterin entwickelter Entscheidungsunterstützung via SaaS gemeinsame!



(1) *Personenbezogene Daten müssen*

- a) *in einer für die betroffene Person nachvollziehbaren Weise verarbeitet werden*  
(„ [...] **Transparenz**“)[.]

- Blackbox-Problematik bei LLMs, insb. beim Einsatz für Entscheidungen
- Mangelnde Transparenz kann menschliche Aufsicht beeinträchtigen (s. auch Art. 14, 13 KI-VO)
- Müssen wir Entscheidungen erklären (zB nach Art. 13 II lit. f DSGVO) ?
  - Wohl **kein Recht auf Erklärung** nach der DSGVO *Wachter/Mittelstadt/Floridi, International Data Privacy Law, Volume 7, Issue 2, May 2017, Pages 76–99, <https://academic.oup.com/idpl/article/7/2/76/3860948>*
  - Involvierte Logik bei LLMs: Statistische Korrelationen

# 03 LLMs im Datenschutzrecht

Transparenz (Art. 5 I lit. a DSGVO) – Neuere Entwicklungen: Reasoning Models

- Reasoning  $\approx$  **Argumentation, logisches Denken**
- LLMs haben **Defizite** dabei (zB beim legal reasoning), weil sie Approximationsmodelle sind und auf statistischen Korrelationen beruhen
- Bessere Ergebnisse lassen sich mit **schrittweiser Herangehensweise** erzielen
- Variante A: **Prompting** zur schrittweisen Vorgehensweise (step-by-step bzw. chain-of-thought-Prompting), das Modell simuliert dann die entsprechende Logik
- Variante B: **Reasoning Models**
  - auf Reasoning-Tasks **feinabgestimmt**
  - ggfs. Implementierung **präziser Regeln, Logiken** (zB symbolic solvers, dh nicht nur Machine-Learning/Approximation) oder **Tools** wie Taschenrechnern
  - Abgrenzung** zu Non-Reasoning Models im Einzelnen unscharf
  - Selbst bei rein statistischer Basis Logik „verschoben“: **Attention** ( $\approx$  Gewichtung Bedeutung von Textbestandteilen in best. Kontext)

I have 72 Apples, I give half of them away, I eat 4, 2 rot and I throw them away, how many have I left?

Let's break it down step by step:

1. Start with: 72 apples
2. Give away half:
  - $72 \div 2 = 36$  given away
  - Remaining =  $72 - 36 = 36$  apples
3. Eat 4 apples:
  - $36 - 4 = 32$  apples left
4. 2 rot and are thrown away:
  - $32 - 2 = 30$  apples left

✅ **Final Answer: 30 apples**

Would you like me to show how a reasoning-trained model might explain this differently or more systematically? 

ChatGPT: Beispiel für schrittweises, mathematisches Reasoning

### (1) Personenbezogene Daten müssen

- a) in einer für die betroffene Person nachvollziehbaren Weise verarbeitet werden  
(„ [...] **Transparenz**“)[.]

- Blackbox-Problematik bei LLMs, insb. beim Einsatz für Entscheidungen
- Mangelnde Transparenz kann menschliche Aufsicht beeinträchtigen (s. auch Art. 14, 13 KI-VO)
- Müssen wir Entscheidungen erklären (zB nach Art. 13 II lit. f DSGVO) ?
  - Wohl **kein Recht auf Erklärung** nach der DSGVO *Wachter/Mittelstadt/Floridi, International Data Privacy Law, Volume 7, Issue 2, May 2017, Pages 76–99, <https://academic.oup.com/idpl/article/7/2/76/3860948>*
  - Involvierte Logik bei LLMs: Statistische Korrelationen
  - Ggfs. beim Chain-of-thought-Prompting/bei Reasoning-Modellen auch Auskunft über die imitierte/nachgebildete Logik bzw. die einzelnen Schritte
- Aber: Für Verwaltung gibt es Begründungserfordernisse  
Dazu *Pesch, European Journal of Risk Regulation (EJRR), Vol. 16 (March 2025), Special Issue 1 (Special Issue on Charting the Landscape of Automation of Regulatory Decision-Making), 76–95. Open Access, <https://www.cambridge.org/core/journals/european-journal-of-risk-regulation/article/potentials-and-challenges-of-large-language-models-llms-in-the-context-of-administrative-decisionmaking/EAE8564B67B49B1D77E948E058935060>*

# 03 LLMs im Datenschutzrecht

Transparenz (Art. 5 I lit. a DSGVO) und Betroffenenrechte

(1) *Personenbezogene Daten müssen*

a) *in einer für die betroffene Person nachvollziehbaren Weise verarbeitet werden*  
(„ [...] **Transparenz**“)[.]

- Ziel: Betroffene sollen wissen, was mit ihren Daten geschieht, insb. **um effektiv Betroffenenrechte** geltend machen zu können (≈ Gewährleistung informationeller Selbstbestimmung i. S. d. deutschen Datenschutzgrundrechts), z. B.
  - Auskunftsrecht, Art. 15 DSGVO
  - Recht auf Vergessenwerden, Art. 17 DSGVO

- Unter Annahme Personenbezug LLM: Berichtigungs- und **Löschungsansprüche in Bezug auf das Modell?**
- Zumindest Behauptung oft: Unlearning/Entfernen **unmöglich; jedenfalls idR sehr aufwändig**
- Problem auch: **Machine Unlearning** bzgl. mancher memorisierter Trainingsdaten macht wahrscheinlicher, dass andere extrahiert werden (**schwächt Datenschutz**) *Hayes et al., arXiv:2403.01218, 2024.*
- Alternativ: **Filterlösungen** ausreichend?

'After colonizing the adult world, AI enters the world of children'

With just a few messages, biased AI chatbots swayed people's political views

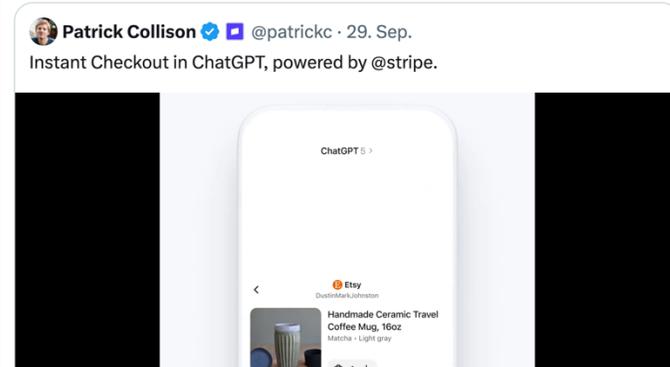
### The Chatbot Culture Wars Are Here

Conservatives, including President Trump, are accusing A.I. companies of left-wing bias, following a playbook that worked well against social media platforms.

Generative AI  
**AI-Generated "Workslop" Is Destroying Productivity**  
by Kate Niederhoffer, Gabriella Rosen Kellerman, Angela Lee, Alex Liebscher, Kristina Rapuano and Jeffrey T. Hancock  
September 22, 2025, Updated September 25, 2025



**richard** @richard\_normal  
these people told you they were going to revolutionize art and science, cure cancer, fundamentally change human existence, and what they have delivered is More Ads  
[Post übersetzen](#)



### AI as Normal Technology

An alternative to the vision of AI as a potential superintelligence

BY **ARVIND NARAYANAN & SAYASH KAPOOR**  
**APRIL 15, 2025**



[paulina.pesch@fau.de](mailto:paulina.pesch@fau.de)



<https://www.linkedin.com/in/paulinajopesch1337/>



<https://bsky.app/profile/paulinajo.bsky.social>



<https://www.digit.rw.fau.de/>

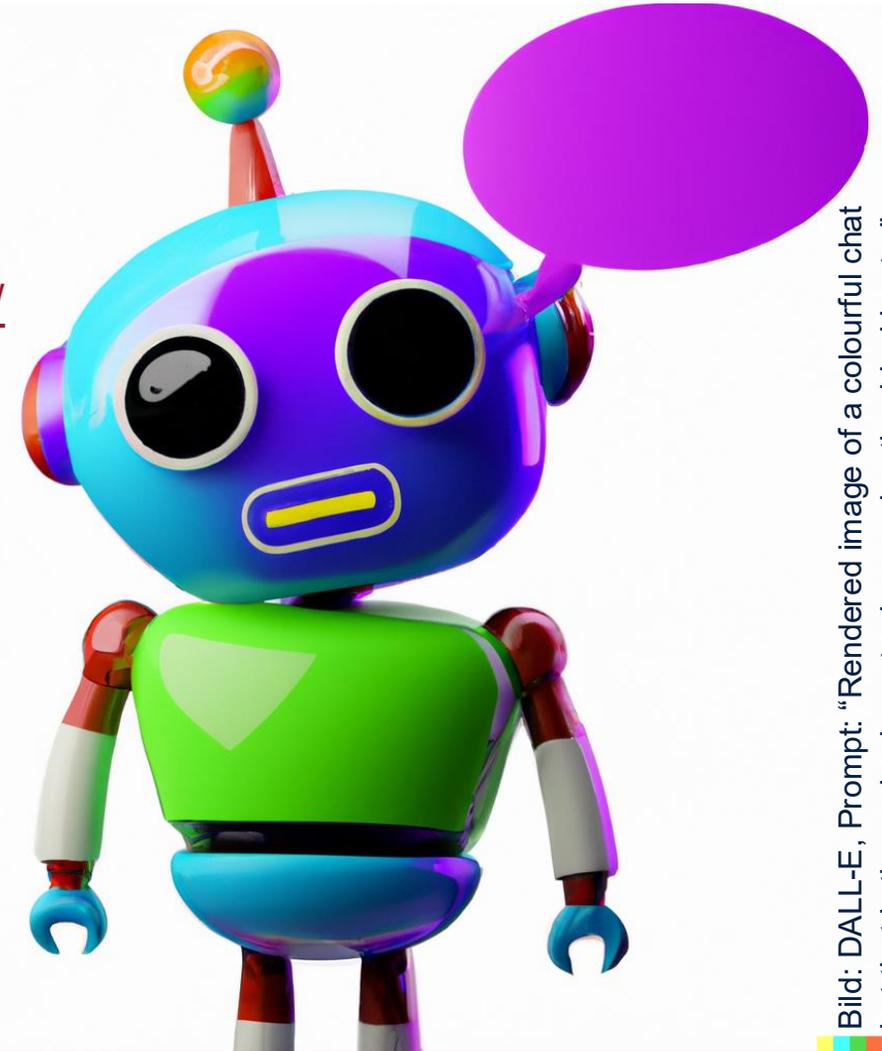


Bild: DALL-E, Prompt: "Rendered image of a colourful chat bot that is the main character in an animation blockbuster".

*In Erscheinung:* Pesch, Paulina Jo, in: Spindler/Schuster/Kaesling, Recht der elektronischen Medien, 5. Aufl. 2025, UrhG, § 44b, Text- und Data-Mining.

Pesch, Paulina Jo/Magnussen, Kristina, LLMs im Forschungsdatenschutz – Rechtmäßigkeit der Nutzung und des Finetunings von LLMs zu wissenschaftlichen Forschungszwecken, DuD 6/2025, Steinrötter (Gasthrsg.), Themenschwerpunkt „KI und Datenschutz“, 367–372.

*Pesch*, Potentials and challenges of Large Language Models (LLMs) in the context of administrative decision-making, European Journal of Risk Regulation (EJRR), January 2025, pp. 1–20. [Open Access](#)

*Kätscher/Pesch*, Interview mit einem Chatbot – Automatisierte Entscheidungsfindung mittels großer Sprachmodelle (LLM) im Beschäftigtenkontext, KIR 2024, 46–55.

*Pesch/Böhme*, Verarbeitung personenbezogener Daten und Datenrichtigkeit bei großen Sprachmodellen (LLM) – ChatGPT & Co. unter der DS-GVO, MMR 2023, 917–923.

*Pesch/Böhme*, Artpocalypse now? Generative KI und die Vervielfältigung von Trainingsbildern, GRUR 2023, 997–1007.

*Pesch*, The case of LAION: The first (German) court decision on text and data mining (TDM) in the context of machine learning – Comment on Regional Court of Hamburg, judgment of 27 September 2024 – 310 O 227/23 [with a machine translation of the judgment], LTZ 2025, 65–77. [PDF](#)