

- openHPI: ChatGPT: Was bedeutet generative KI für unsere Gesellschaft? -

# Zusammenfassung Woche 1

Johannes Hötter  
Christian Warmuth

# Vor allem interessant

- Kreative Berufe & Wissensarbeit wird "automatisiert"
- Sprache als etwas Grundmenschliches & bisher sehr schwer zu meistern
- Chat-Interface als vielseitige Schnittstelle
- Fokus auf Sprache als Mittel der Interaktion

## **Zusammenfassung Woche 1**

Johannes Hötter,  
Christian Warmuth

openHPI

Folie 2

# Diskriminative vs. generative KI

## Diskriminativ

Klassifikation

Extraktion

Suche

- Identifiziert Muster oder Beziehungen in vorhandenen Daten
- Konzentriert sich auf die Klassifizierung von Daten
- Wird verwendet, um Entscheidungen oder Vorhersagen zu treffen

## Generativ

Generierung

Personalisierung

- Fokus auf Generierung und Personalisierung von Inhalten
- z.B. "Dieser Text ist zu komplex, fasse die wesentlichen Aussagen zusammen"
- Kombination mit diskriminativer KI möglich!

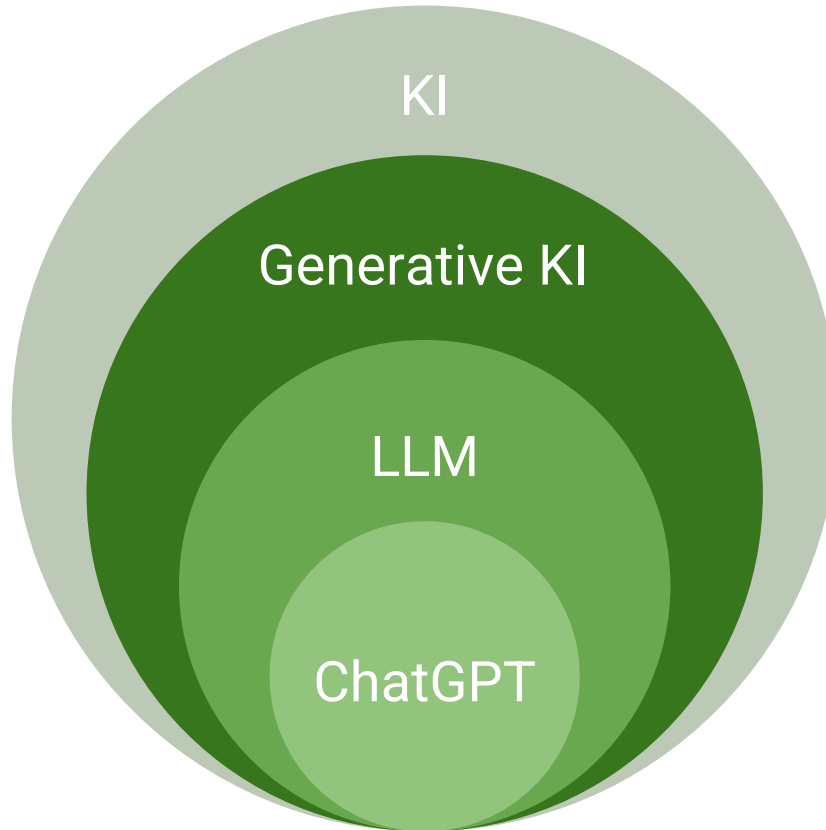
## Zusammenfassung Woche 1

Johannes Hötter,  
Christian Warmuth

openHPI

Folie 3

# Wichtige Abgrenzung



## Zusammenfassung Woche 1

Johannes Hötter,  
Christian Warmuth

openHPI

# Generative KI



**Bild-Modelle**  
(Diffusion-Modelle)

... again, Mummie, that wouldn't be any railway fare and he shouldn't  
oms. Oh, do let us go in a caravan."  
Mrs. Russell shook her head. "I know it sounds lovely, darling; but  
; we to get a caravan? It would cost at least fifty pounds to buy one  
en if we had one, Daddy couldn't get away this summer. No, we  
ike up our minds to do without a holiday this year; but I'll tell you wh  
ll do: we'll all go to Southend for the day, as we did last year, and  
r lunch and tea with us and have a splendid picnic."  
"Then we can bathe again," said Bob; "but, oh! I do wish I could h  
ny and ride," he added unexpectedly. "You don't know how I long  
ny," he continued, sighing deeply as he remembered the blissful ho  
en a friend let him share his little Dartmoor pony and ride occasiona  
"Southend is nothing but houses and people," cried Phyllis; "it's no l  
an this place; and oh! Mummie, I do so *long* for fields and flower  
imals," she added piteously; and she shook her long brown hair fo  
hide the tears in her eyes.  
"Never mind, darling, you shall have them one day," answered  
assell with easy vagueness.  
This really was not very comforting, and it was the most fortunate thing  
at at that moment a car stopped at the door.  
"Uncle Edward!" shouted Bob, rushing from the room. Phyllis br  
e tears so hastily from her eyes that she arrived at the front door alm  
on as he did, and both flung themselves on the tall, kindly-looking man s  
g beside the car.  
"Uncle Edward! Uncle Edward!" they cried. "You've come at  
e've been longing to see you. Oh, how glad we are you're here!"  
Now the delightful thing was that their uncle seemed just as pleased  
em as they were to see him, and returned their hugs and greetings wit  
most cordiality. They were just on the point of dragging him int  
ouse, hanging one on each arm, when he said: "Stop, not so fast. The  
me things to fetch in from the car."  
So saying he began diving into the back of it and bringing out, not c  
itcase, but various parcels, which he handed out one by one.  
"That's the pair of chickens Pops brought for your mother," said he, to

**LLMs**

...

## **Zusammenfassung Woche 1**

Johannes Hötter,  
Christian Warmuth

openHPI

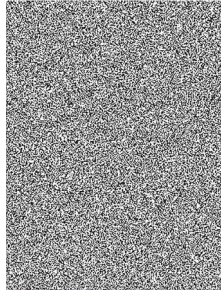
Folie **5**

# Wie entstehen Bilder mit generativer KI?



Person wirft eine Kamera in die Luft.

Encoding



Person wirft eine Kamera in die Luft.

Decoding



Person wirft eine Kamera in die Luft.

## Zusammenfassung Woche 1

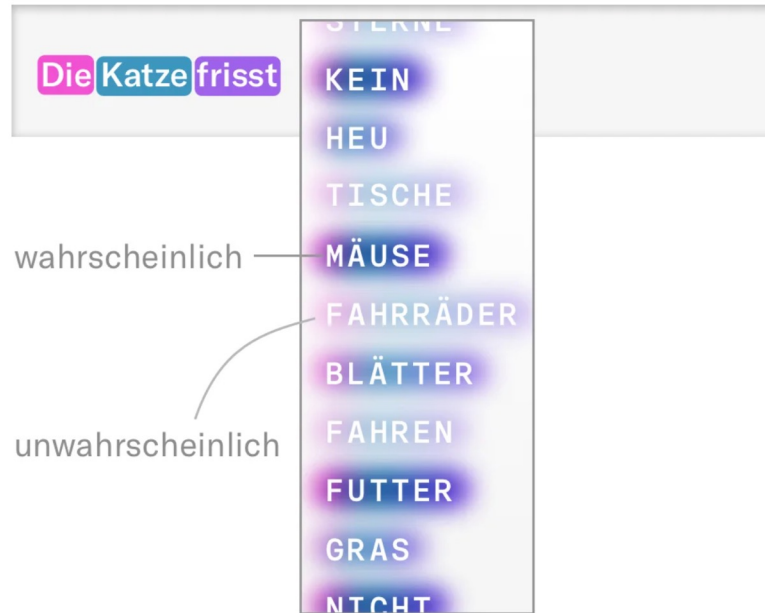
Johannes Hötter,  
Christian Warmuth

openHPI

Folie 6



# Was macht ein Sprachmodell?



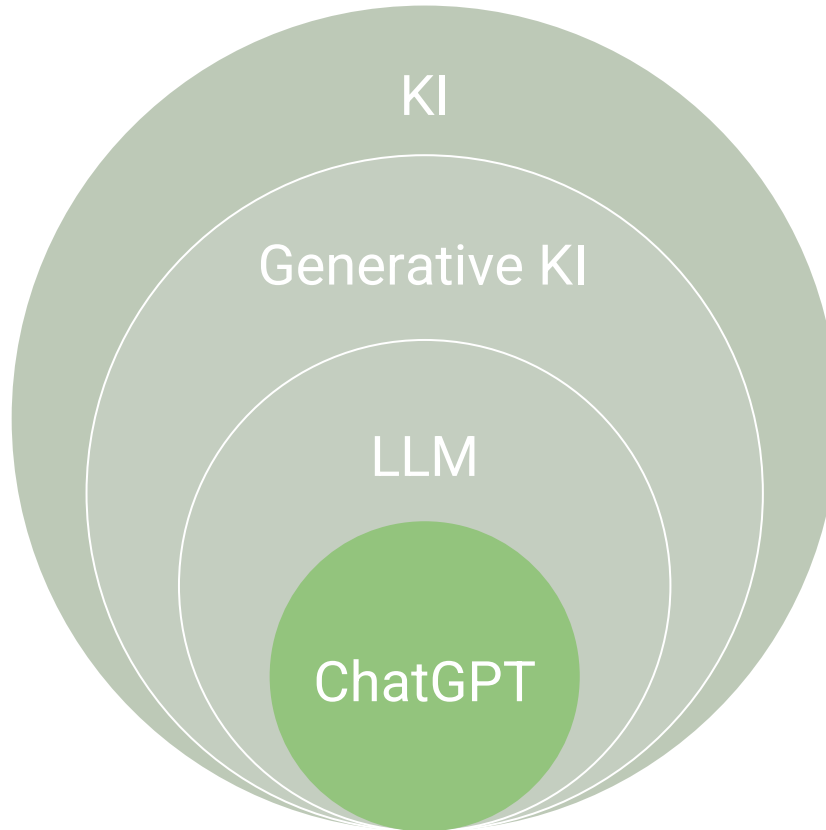
Quelle: NZZ

## Zusammenfassung Woche 1

Johannes Hötter,  
Christian Warmuth

openHPI

# Wie funktioniert ChatGPT?



## **Zusammenfassung Woche 1**

Johannes Hötter,  
Christian Warmuth

openHPI



# Funktionsweise von ChatGPT

**1. Pretraining des Large Language Models**  
(99% Trainingszeit & Kosten)

**2. Fine-Tuning um Konversationen führen zu können**

**3. Reward Modell das menschliche Präferenzen abbildet**

**4. Kontinuierliches Reinforcement Learning**  
(basierend auf menschlichem Feedback)

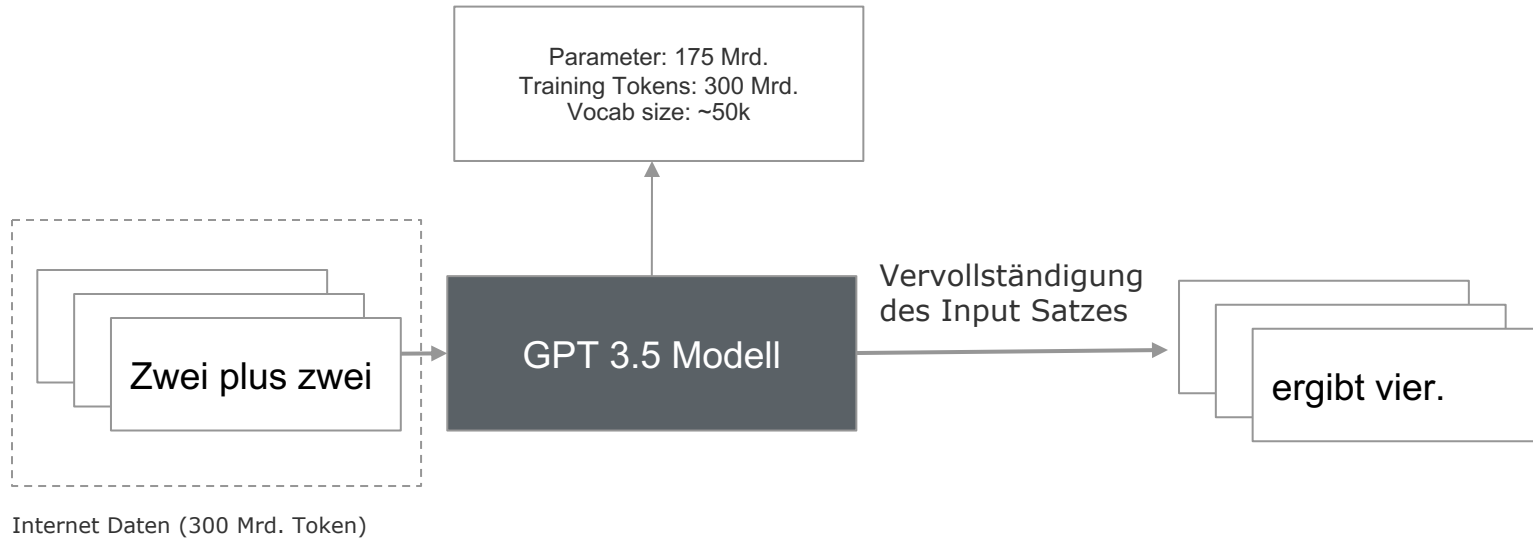
**Zusammenfassung  
Woche 1**

Johannes Hötter,  
Christian Warmuth

openHPI

Folie **9**

# Schritt 1: Pretraining des LLM



**Zusammenfassung  
Woche 1**

Johannes Hötter,  
Christian Warmuth

openHPI

Folie **10**

## Schritt 2 + 3:

### Konversationsbeispiele sammeln & Reward Modell bauen



Was mache ich, wenn es regnet?

**- 4 Punkte**

Ich springe von links nach rechts auf meinen Fersen.

**+5 Punkte**

Ich spanne meinen Regenschirm auf.

**- 2 Punkte**

Ich starte meinen Laptop.

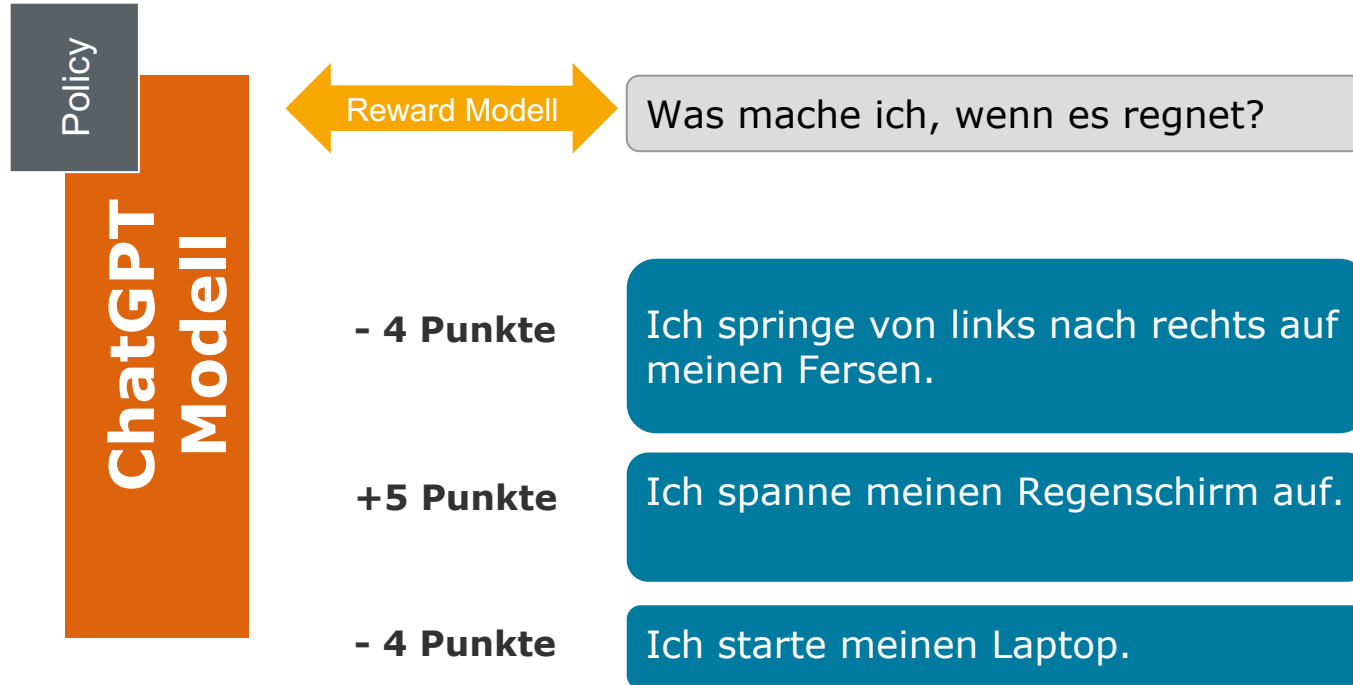
#### **Zusammenfassung Woche 1**

Johannes Hötter,  
Christian Warmuth

openHPI

Folie **11**

# Schritt 4: Reinforcement Learning on Human Feedback



## Zusammenfassung Woche 1

Johannes Hötter,  
Christian Warmuth

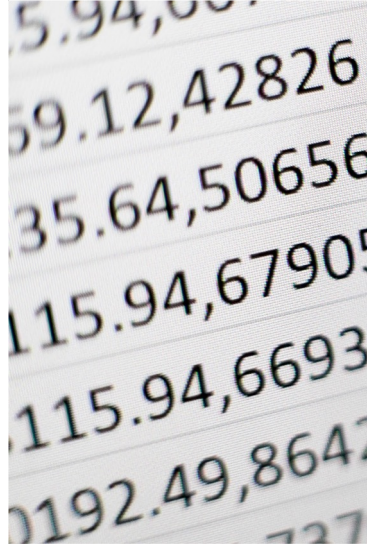
openHPI

Folie **12**

# Was sind Stellschrauben für gute Ergebnisse? (wenn man nicht selbst LLMs von Grund auf trainiert)



Modell-  
Verbesserung



Fine-Tuning auf  
eigenen Daten



Prompt  
Engineering

## **Zusammenfassung Woche 1**

Johannes Hötter,  
Christian Warmuth

openHPI

Folie 13

# Verschiedene Wege der Umsetzung generativer KI

## Prompt Engineering

### Kontext mitliefern

**Informationen** werden im Kontext (Prompt) mitgeben um bessere Ergebnisse zu erhalten (z.B. via Vektordatenbank)

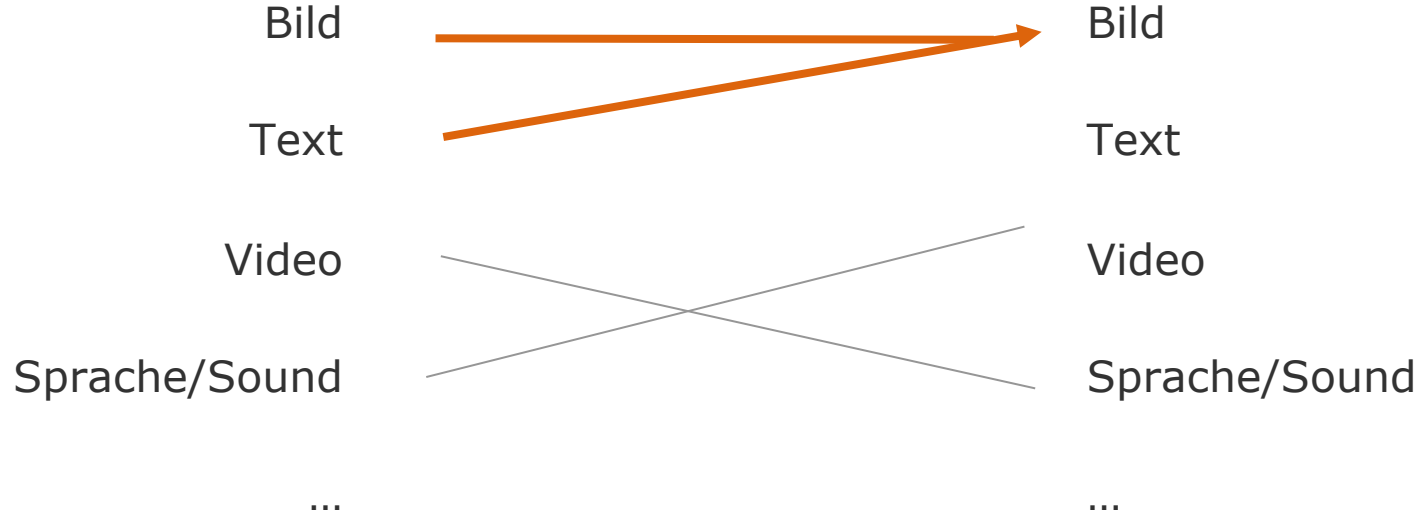
### In-Context Learning

**Beispiele** im Kontext (Prompt) mitgeben, um bessere Ergebnisse zu erhalten.  
Synonym: **Few-Shot Prompting**

### Fine Tuning

**Beispiele** werden verwendet für das Finetuning. Das Modell wird verändert und verbessert.  
Wichtig: Rechenaufwand bzw. Kosten entstehen.

# Multimodalität

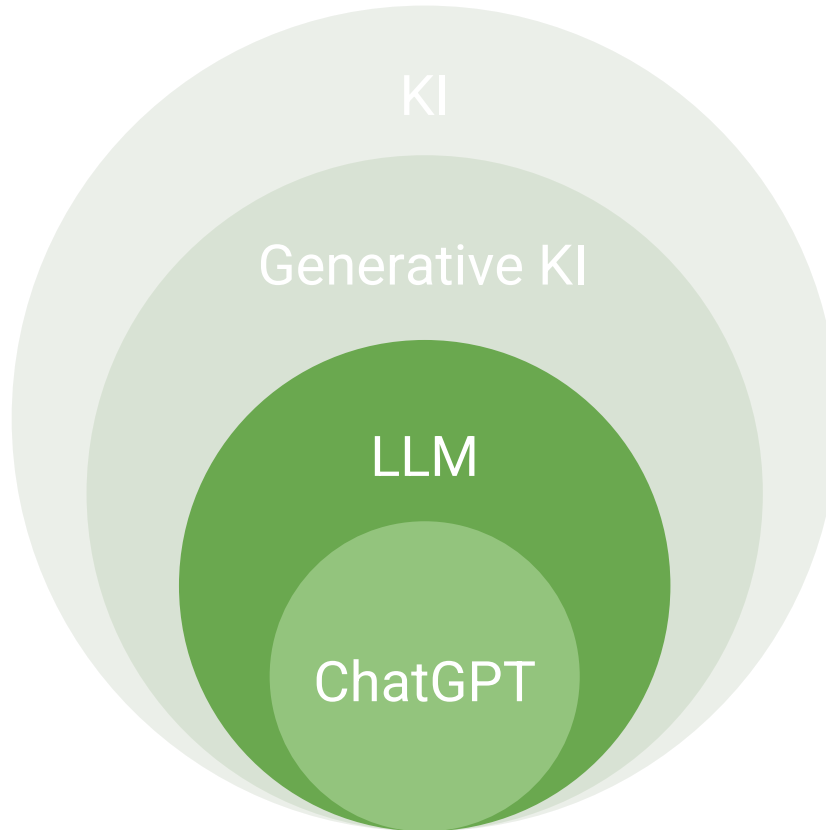


## **Zusammenfassung Woche 1**

Johannes Hötter,  
Christian Warmuth

openHPI





## **Zusammenfassung Woche 1**

Johannes Hötter,  
Christian Warmuth

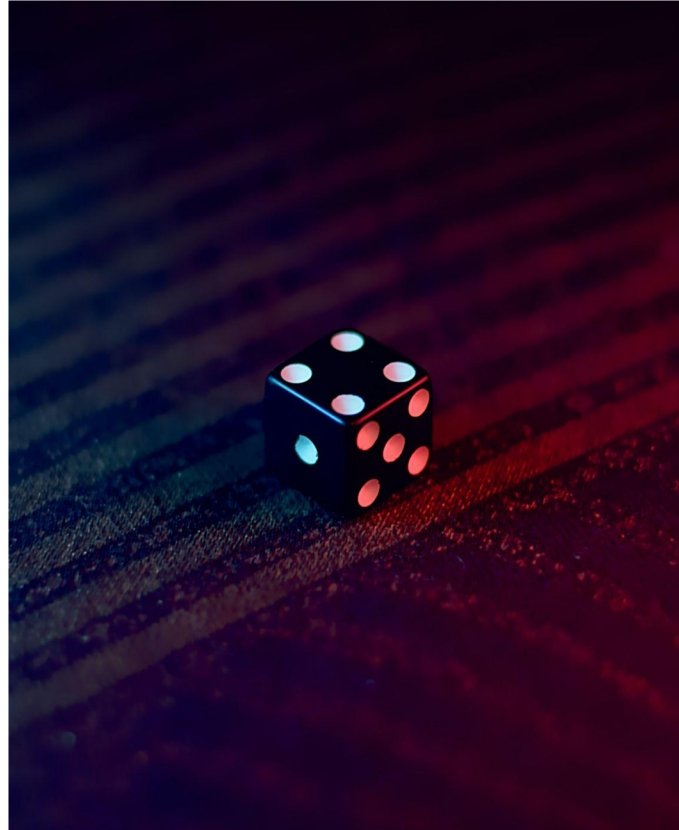
openHPI

# Generative KI-Modelle sind nicht deterministisch!

Ergebnisse von generativen KI Modellen sind nicht deterministisch!

Das heißt, dass man bei gleicher Eingabe verschiedene Ergebnisse erhalten kann.

(Außer man legt einen sogenannten "Seed" Parameter fest).



## **Zusammenfassung Woche 1**

Johannes Hötter,  
Christian Warmuth

openHPI

Folie **17**

- openHPI: ChatGPT: Was bedeutet generative KI für unsere Gesellschaft? -

# Zusammenfassung Woche 1

Johannes Hötter  
Christian Warmuth