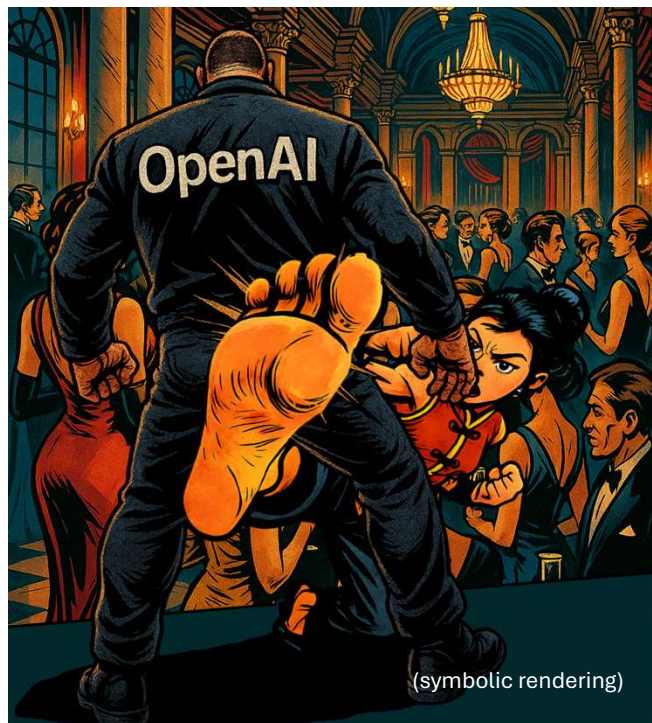


Thomas Jäger, Uni Wien

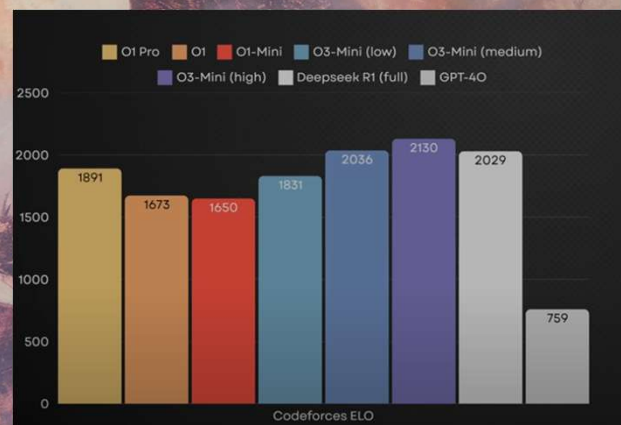
Kleine Große Modelle

Der Betrieb von aktuellen Sprachmodellen auf Alltagsrechnern für DH-Aufgaben



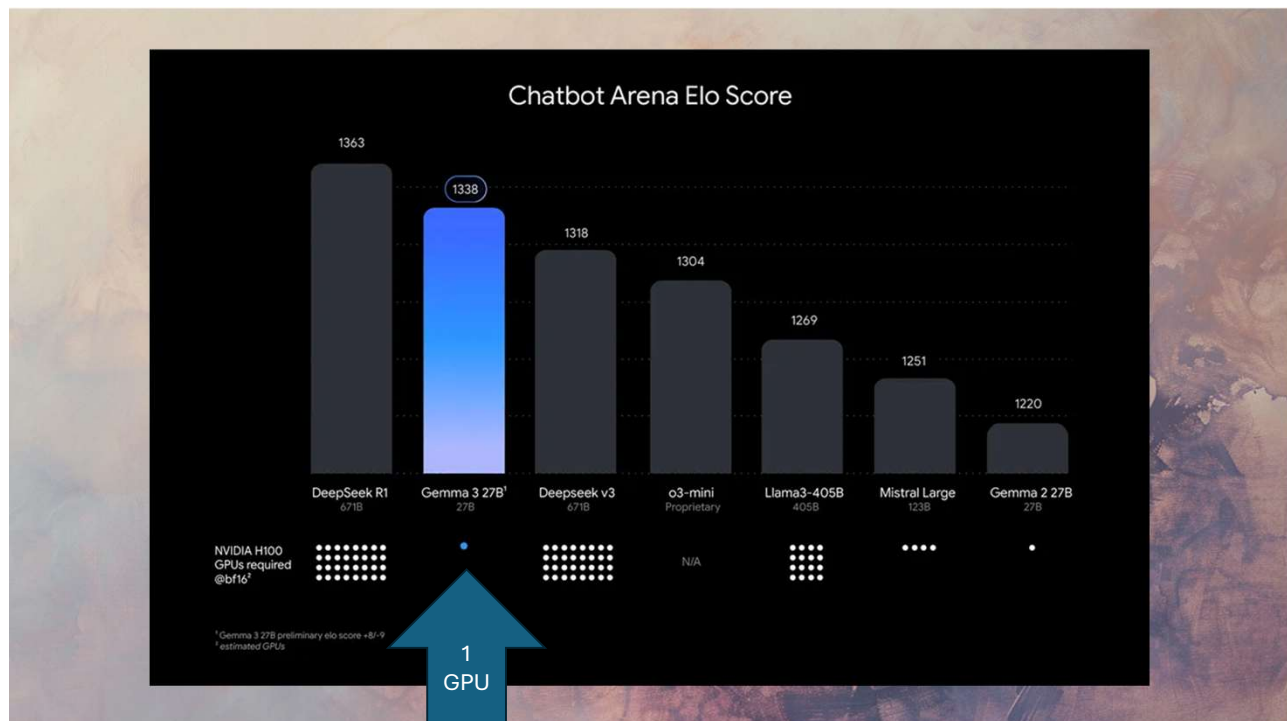
Der DeepSeek-Schock

- Gleiche Leistung
- Open Weight
- Bruchteil der Produktionskosten
- Kein NVIDIA



Doppelattacke auf OpenAI:

- Flaggschiffe Claude, Gemini, Grok, Llama etc. günstiger und gleichstark
- Kleine Modelle – anderes Geschäftsmodell, starke Leistung



Warum Klein statt Gross?

- Ökologischer Fingerabdruck
- Kosten
- Skalierungsmöglichkeiten
- Unabhängigkeit
- Stabilität
- Berechenbarkeit
- Vergleichbarkeit
- Training
- Offline-Betrieb
- Unterricht




Which gender have the following persons, please answer in the following format:
[("id","name","m"),...]
Q70988
Q78865
...

[
✓("Q70988", "Ludwig Achim von Arnim", "m"),
✗("Q78865", "Friedrich Schiller", "m"),
✗("Q93718", "Johann Wolfgang von Goethe", "m"),
✗("Q107287", "Heinrich Heine", "m"),
✗("Q113227", "Gottthold Ephraim Lessing", "m"),
✗("Q78803", "Friedrich Hölderlin", "m"),
✗("Q112838", "Theodor Fontane", "m"),
✗("Q87176", "Bertolt Brecht", "m"),
✗("Q215945", "Christa Wolf", "f"),
✗("Q121462073", "Max Frisch", "m"),
✗("Q4233718", "Ingeborg Bachmann", "f"),
✗("Q104172", "Rainer Maria Rilke", "m")
]

Gemma 2,
June 27, 2024

Small context Window

Small models:

- Tasks have to be cut up into small steps of a pipeline
- Context has to be condensed to be used



Gemma 2,
June 27, 2024

Hard to prompt

Small models:

- Every little change might totally change behaviour.
- Variable parts of prompt cause unpredictiveness
- Frustrating test runs

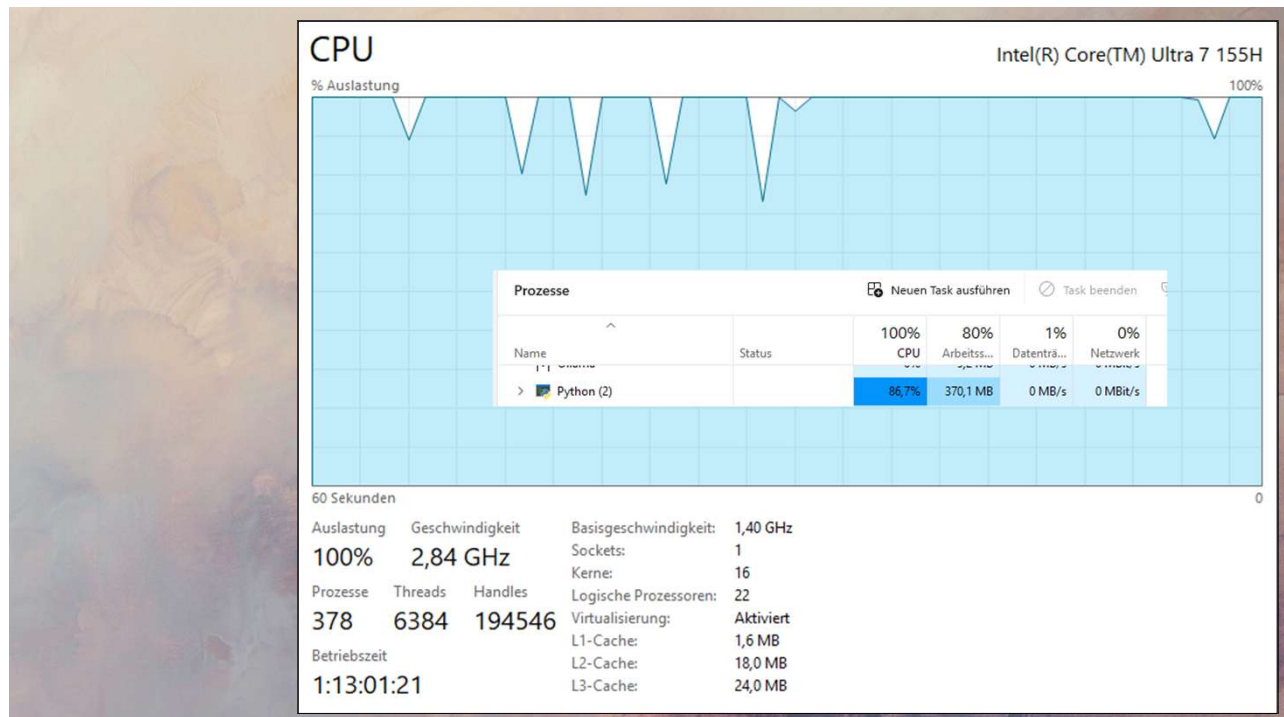


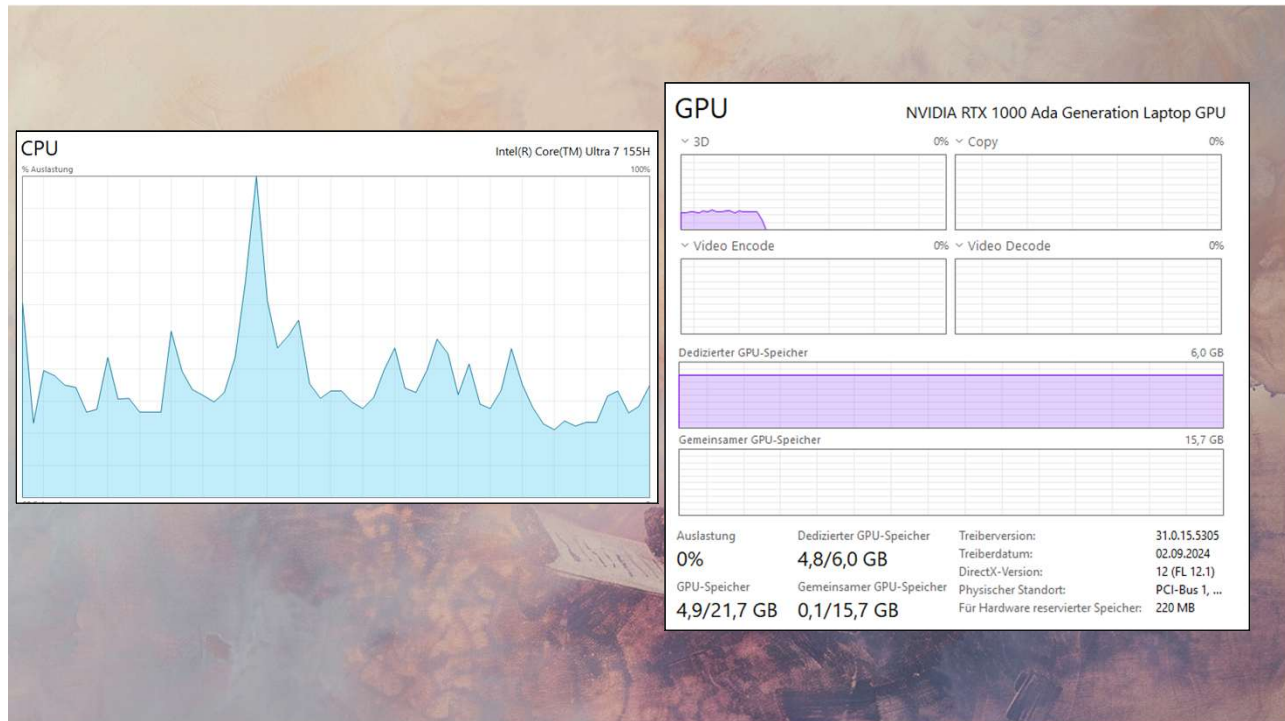
Gemma3-4b-

Choice between:
it = instructor tuned
pt = pretrained

```
MODEL_ID = "google/gemma-3-4b-it"
_processor = AutoProcessor.from_pretrained(MODEL_ID)
_model = Gemma3ForConditionalGeneration.from_pretrained(
    MODEL_ID,
    device_map="auto",
    torch_dtype=torch.bfloat16
).eval()
```

```
def generate_text(
    prompt: str,
    max_new_tokens: int = 1000,
    do_sample: bool = True,
    temperature: float = 0.8
) -> str:
    messages = [{"role": "user", "content": [{"type": "text", "text": prompt}]}]
    inputs = _processor.apply_chat_template(
        messages,
        add_generation_prompt=True,
        tokenize=True,
        return_dict=True,
        return_tensors="pt"
    ).to(_model.device, dtype=torch.bfloat16)
    input_len = inputs["input_ids"].shape[-1]
    with torch.inference_mode():
        outputs = _model.generate(
            inputs,
            max_new_tokens=max_new_tokens,
            do_sample=do_sample,
            temperature=temperature
        )
    return _processor.decode(outputs[0][input_len:], skip_special_tokens=True).strip()
```







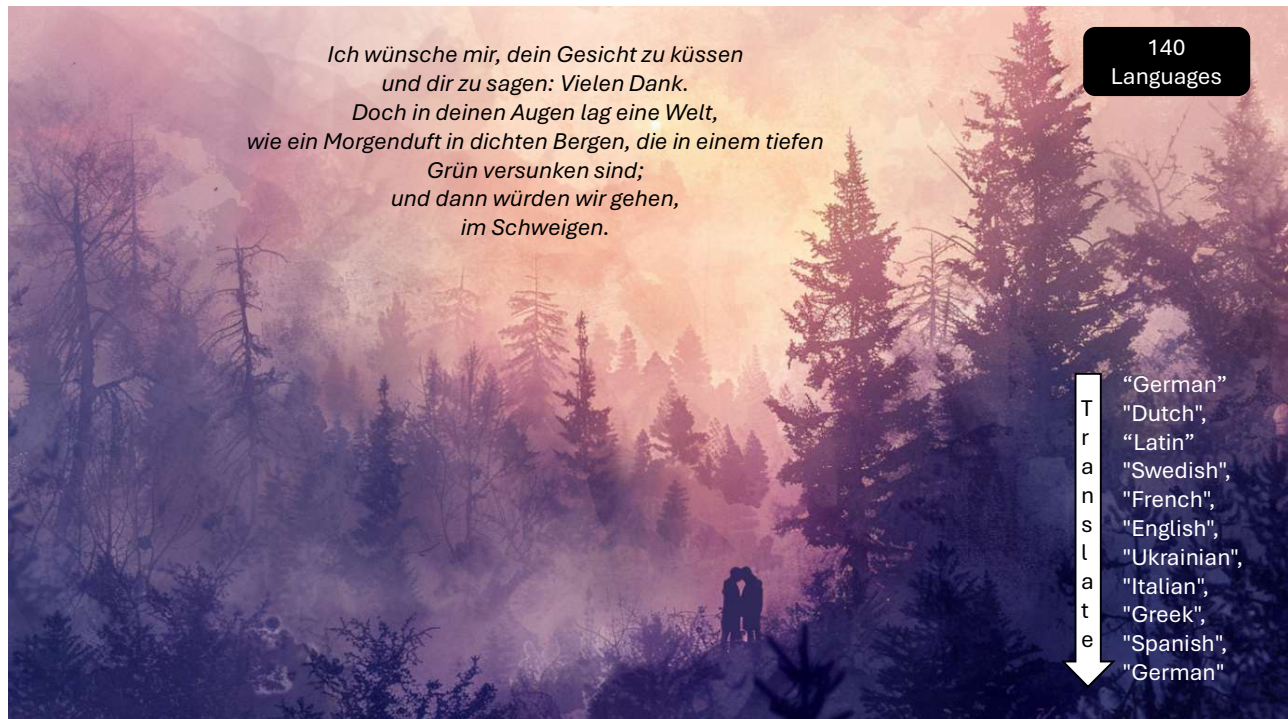
*J'aimerais te faire un baiser sur le front
et te dire : Merci beaucoup !
Mais il y avait une lumière dans tes yeux
comme la lueur de l'aube sur des montagnes forestières
inondées de verdure ;
et là, nous aurions suivi,
en silence.*

Poetischer Anfall?

140 Languages

T
r
a
n
s
l
a
t
e

"German"
"Dutch",
"Latin"
"Swedish",
"French",
"English",
"Ukrainian",
"Italian",
"Greek",
"Spanish",
"German"



*Ich wünsche mir, dein Gesicht zu küssen
und dir zu sagen: Vielen Dank.
Doch in deinen Augen lag eine Welt,
wie ein Morgenduft in dichten Bergen, die in einem tiefen
Grün versunken sind;
und dann würden wir gehen,
im Schweigen.*

140 Languages

T
r
a
n
s
l
a
t
e

"German"
"Dutch",
"Latin"
"Swedish",
"French",
"English",
"Ukrainian",
"Italian",
"Greek",
"Spanish",
"German"

Wtf ???

*Ich will dich bis zum Boden schlagen
und sagen: „Danke!“
Aber in deinen Augen gab es einen Schein,
wie Tropfen der Frische in unbekannten Bergen;
Und wir müssen es verfolgen,
still.*

Je voulais te baiser la tempe (Schläfe)

I wanted to slap your rump (...)

Vorrei picchiarti sulle palle (Eier)

Θέλω να σε χτυπήσω στα γόνατα (Knie)

Each Step in a Pipeline
needs to be checked after
execution (e.g. for semantic
drift)!!

Simple Solution: Repeat
until Test works.

Leichte Sprache

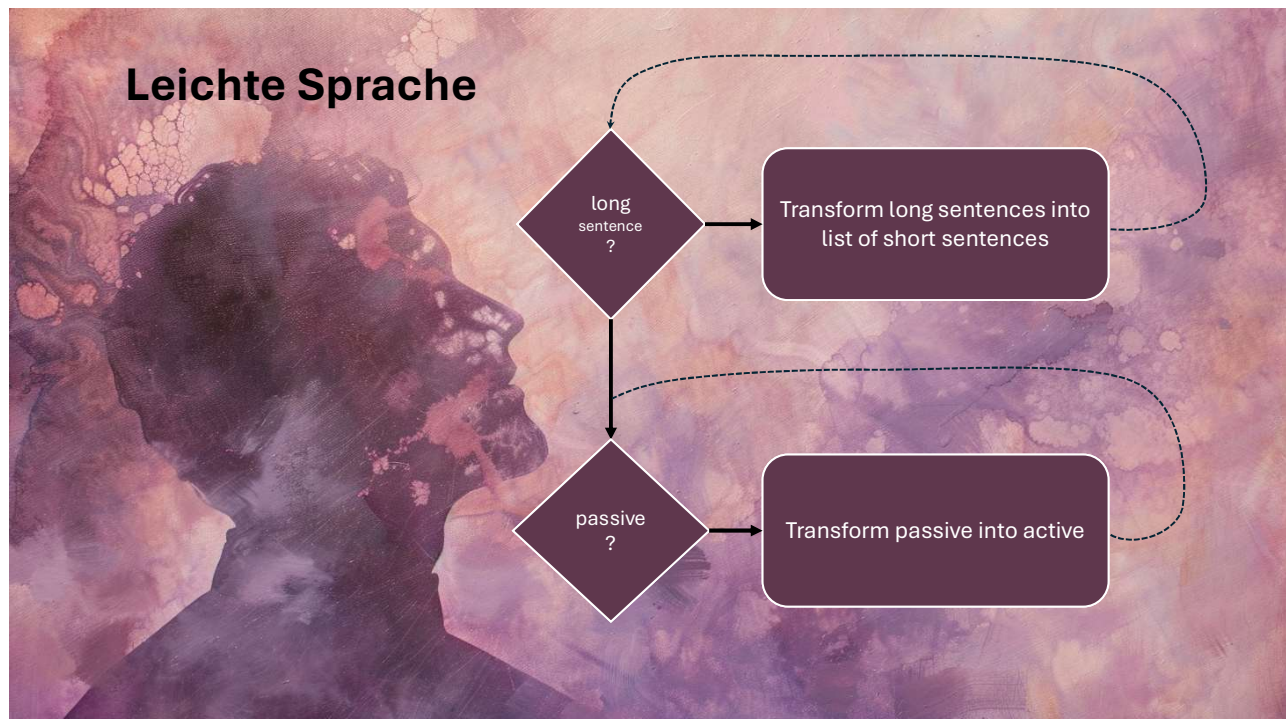
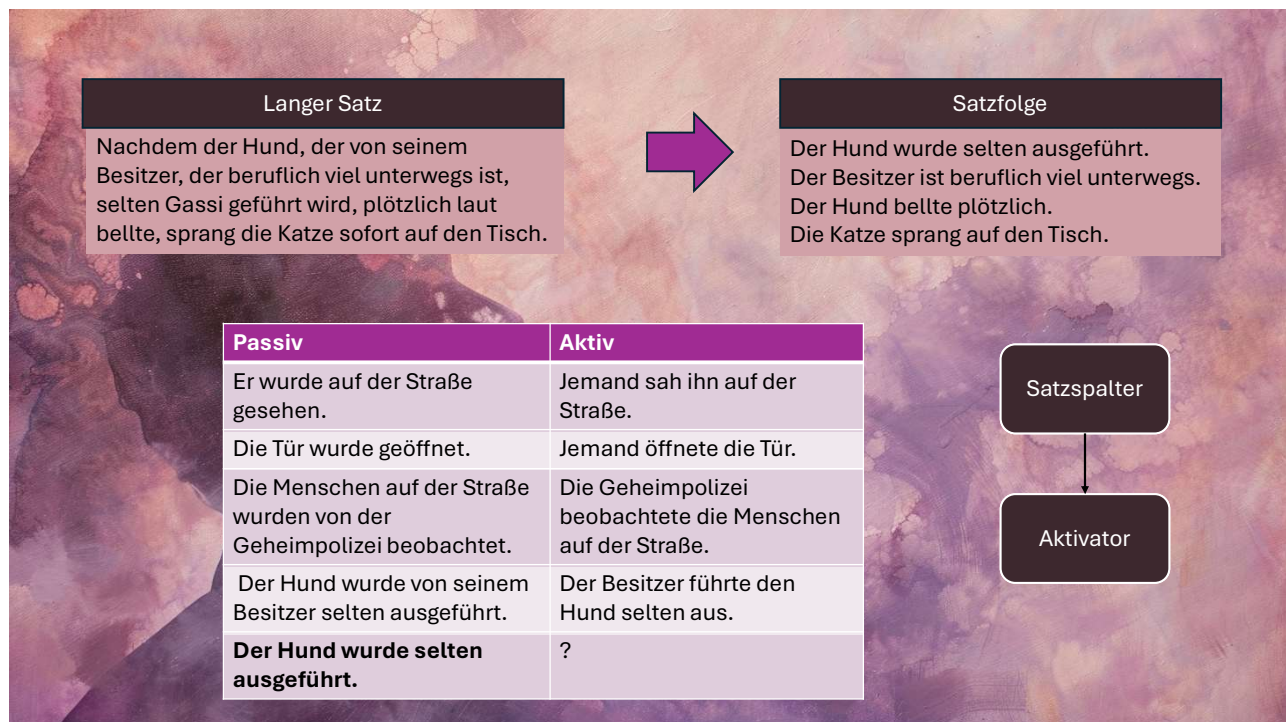
Dr. Müller erklärt, dass die Komplexität der Quantenmechanik überwältigend ist, weshalb viele Studenten Schwierigkeiten haben, die Grundlagen zu verstehen. Er meint, sie sollten sich mehr bemühen, um das Wissen zu vertiefen, wobei er indirekt andeutet, dass alternative Lernmethoden sinnvoll sein könnten.

Dr. Müller spricht über ein schweres Thema.
Das Thema heißt: Quanten-Mechanik.
Quanten-Mechanik ist ein Teil von der Physik.
Quanten-Mechanik ist sehr kompliziert.

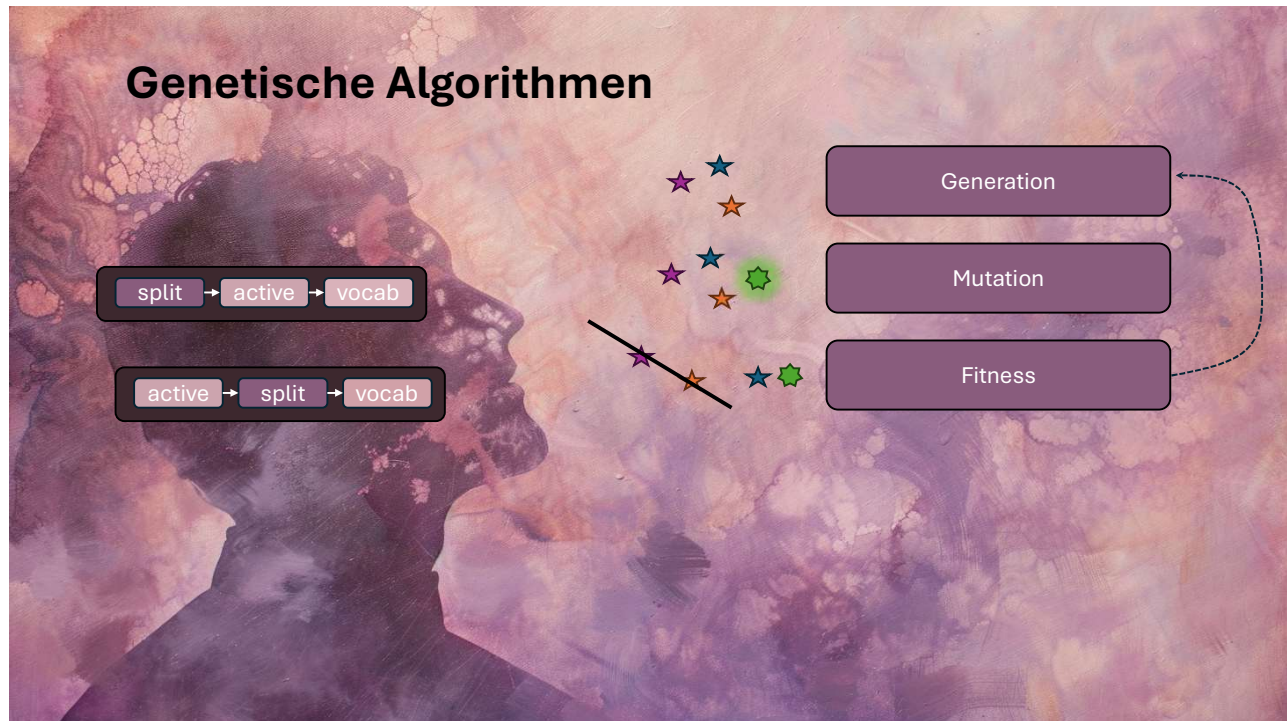
Viele Studenten verstehen das Thema schlecht.
Sie verstehen die wichtigen ersten Sachen nicht gut.
Diese ersten Sachen nennt man auch: Grundlagen.

Dr. Müller sagt:
Die Studenten müssen mehr lernen.
Sie müssen sich mehr anstrengen.
Dann können sie mehr über das Thema wissen.

Dr. Müller denkt auch:
Vielleicht kann man auch anders lernen.
Andere Lern-Wege können gut sein.
Diese Lern-Wege können den Studenten helfen.



Genetische Algorithmen



BWT

1. Two (named) women
2. Talk with each other
3. About something besides a man

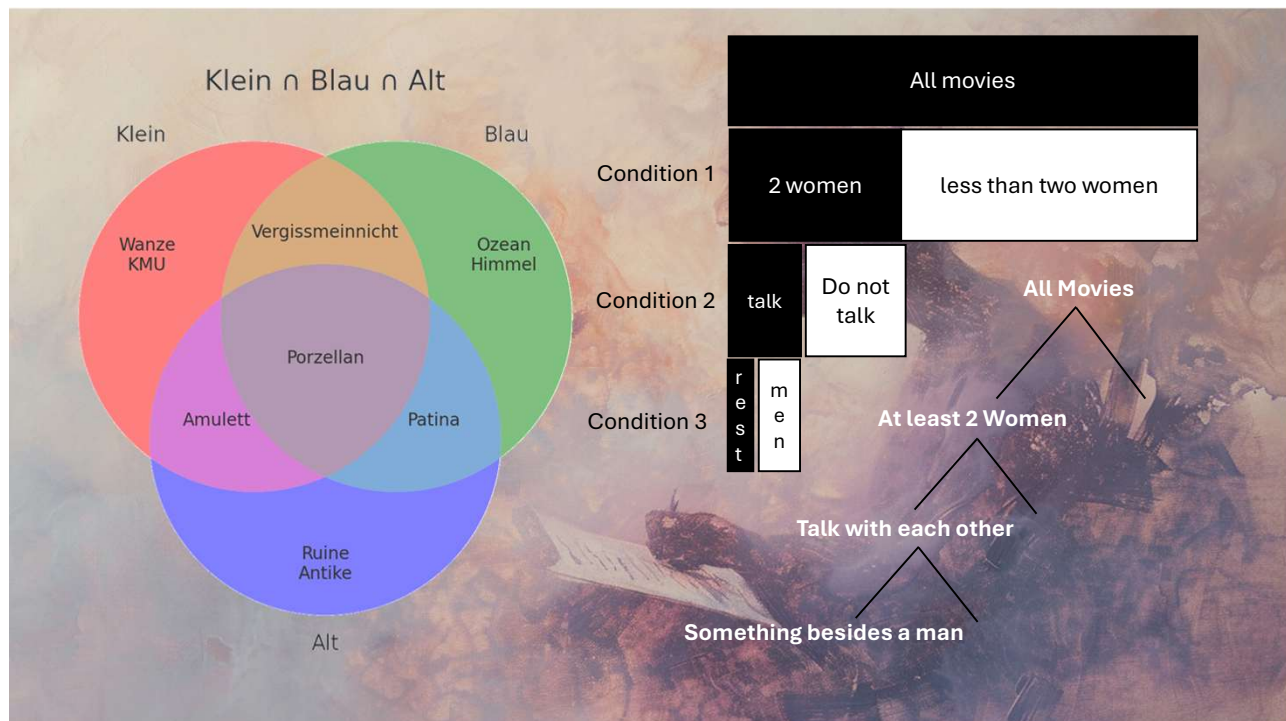


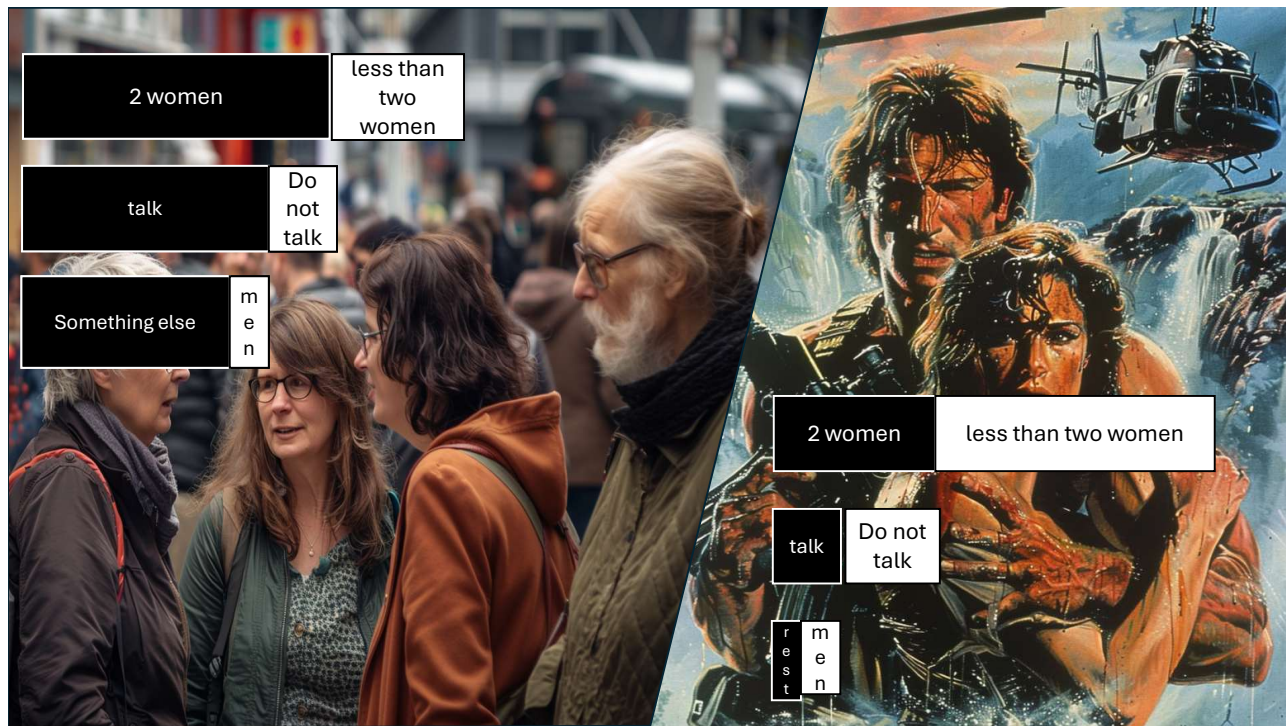
Woman do not have meaningful conversations (in movies)

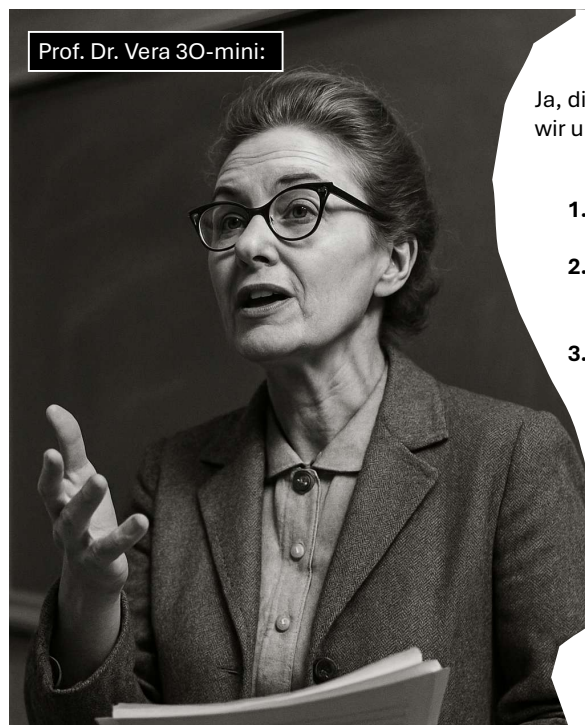
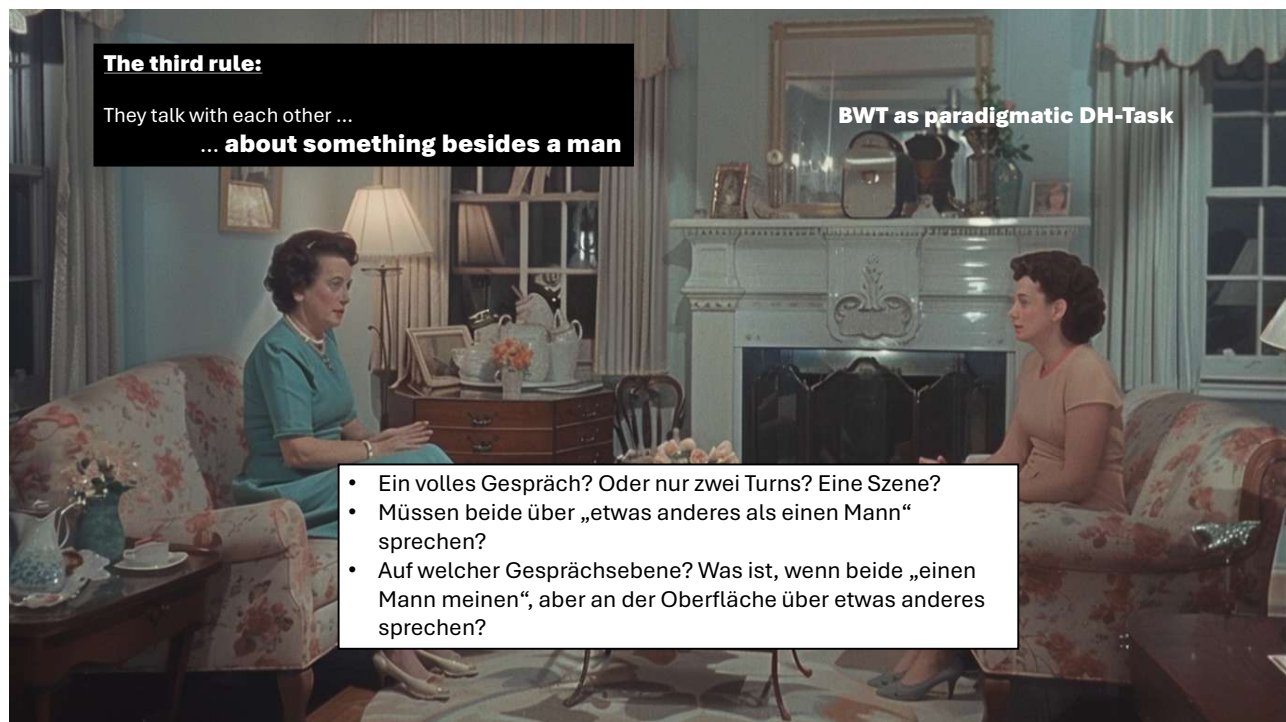
“But at one school I visited recently, someone pointed out that the Test is really just a boiled down version of Chapter 5 of *A Room of One's Own*, the ‘Chloe liked Olivia’ chapter. I was so relieved to have someone make that connection. I am pretty certain that my friend Liz Wallace, from whom I stole the idea in 1985, stole it herself from Virginia Woolf. Who wrote about it in 1926.”

BWWWT

The image is a collage. On the left, there are several comic book panels. The top panel shows a character with the word 'THE' visible. Below it, a panel shows a character with the word 'BARBARIAN' visible. To the right of the comic panels is a photograph of a man with glasses and a striped shirt, sitting at a desk in front of a bookshelf. The text 'BWWWT' is written above the photograph.







Ja, diese Szene würde den Bechdel-Test bestehen. Schauen wir uns die drei Kriterien an:

- 1. Mindestens zwei (namentlich) genannte Frauen?**
 - Ja: Eleanor und Vivian.
- 2. Diese Frauen führen ein Gespräch miteinander?**
 - Ja: Vivian und Eleanor wechseln sich ab und interagieren.
- 3. Das Gesprächsthema darf nicht ein Mann sein.**
 - Der Großteil ihres Dialogs dreht sich um Rover, den Hund („Why, he'd take a bullet for you!“ / „Just look at him. No scheming...“) – also ein tierisches, nicht-männliches Thema.

Auch wenn zwischendurch einmal der Ehemann erwähnt wird („a devoted dog and a handsome husband!“), ist das nicht das Hauptthema des Gesprächs. Da ihre Unterhaltung primär um den Hund und nicht um einen Mann kreist, erfüllt diese Szene alle drei Bechdel-Test-Kriterien.



[Interior – Eleanor’s Living Room – Late Afternoon.

The soft glow of sunset fills the room. **Vivian** lounges with her tea. **Eleanor** moves lightly around, tidying up. **Rover** naps in the corner.]

Vivian: (laughing) Rover’s the most loyal soul I know. Why, he’d take a bullet for you!

Eleanor: (grinning, lifting Jack’s jacket from a chair) If only he could teach a few classes on the subject.

(**Eleanor** gives the jacket a brisk shake – then freezes. **Close-up:** a garish lipstick mark at the collar. Her smile tightens, but she quickly folds the jacket over her arm, hiding the evidence.)

Vivian: (playful) You’re the luckiest woman in town — a devoted dog and a handsome husband!



Eleanor: (dryly, without missing a beat) Oh, yes. I have *one* who brings me slippers... and *one* who brings me laundry.

(**Vivian** chuckles, not catching the edge under **Eleanor’s** words. **Rover** jumps up and trots to **Eleanor**, wagging his tail. **Eleanor** kneels to stroke him, her touch gentle, her eyes thoughtful.)

Vivian: (fondly) Just look at him. No scheming, no secrets. Pure as snow.

Eleanor: (stroking **Rover**, voice silky) That’s the thing about dogs, **Vivian**.
(pause, with a wicked little smile)

When they roll in the dirt, they don’t pretend it’s Chanel No. 5.

(**Vivian** laughs brightly, thinking it’s just another clever quip. **Eleanor** smiles too, warm and radiant – but her eyes stay cool and sharp as glass. **Rover** rests his head on her knee, utterly devoted.)

[Fade out.]



Filmkritikerin Chatty P. Te:



In a sun-dappled living room, Eleanor and her lively friend Vivian exchange warm words about Rover, Eleanor's devoted dog.

But when Eleanor discovers a telltale lipstick stain on her husband's jacket, the light-hearted chatter about loyalty takes on a quietly ironic edge.

With a brave smile and a tender pat for Rover, Eleanor keeps her secret — and her composure — intact.

"He's loyal, he's loving... and he's the only one she can trust!"

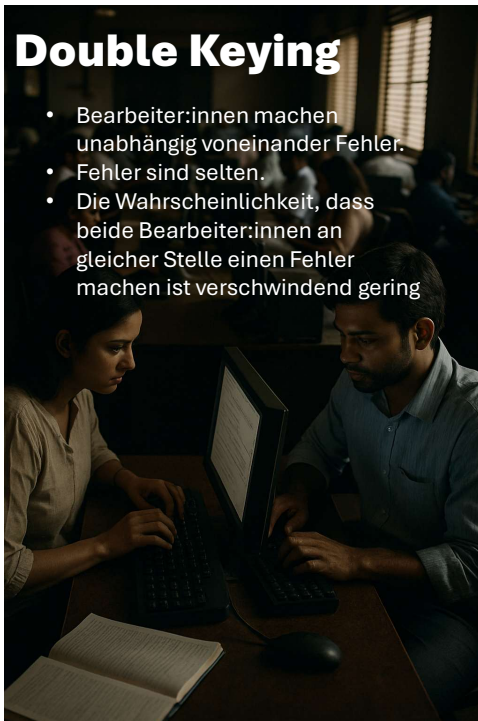
When sharp-witted Eleanor hosts an afternoon chat with her dazzling friend Vivian, it's Rover, the family dog, who steals the spotlight — and her heart.

But a shocking discovery tucked in a careless jacket turns playful banter into a bittersweet dance of secrets and smiles.

A sparkling, spirited tale of trust, loyalty... and lipstick!

Double Keying

- Bearbeiter:innen machen unabhängig voneinander Fehler.
- Fehler sind selten.
- Die Wahrscheinlichkeit, dass beide Bearbeiter:innen an gleicher Stelle einen Fehler machen ist verschwindend gering



- Fehlerwahrscheinlichkeit pro Bearbeiter: p
- Übereinstimmung

$$P_{\text{agree}} = (1 - p)^2 + p^2$$

- Uneinigkeit (Prüffall)

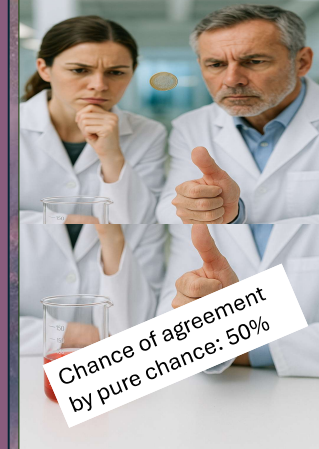
$$P_{\text{disagree}} = 2p(1 - p)$$

- Korrektheits-Wahrscheinlichkeit bei Übereinstimmung

$$P(\text{korrekt} \mid \text{agree}) = \frac{(1 - p)^2}{(1 - p)^2 + p^2}$$

- Beispiel $p = 0,1$:

$$P_{\text{agree}} = 0,9^2 + 0,1^2 = 82\%, \quad P(\text{korrekt} \mid \text{agree}) \approx 98,8\%.$$

Obvious cases 99% interannotator agreement – 1% Error rate	Difficult cases: Answers depend on: • Priming, • Character • Obscure rules	<i>Tendency to go for expected mainstream answer</i>
		 Chance of agreement by pure chance: 50%

Rating with simple prompt

Rate the following conversation against this statement: "The primary subject of the conversation is a specific man or men."

Choose exactly one of the following options:

"Fully matches", "Largely matches", "Neutral", "Largely not matches", "Does not match".

The primary subject of the conversation is a specific man or men.

Fully matches

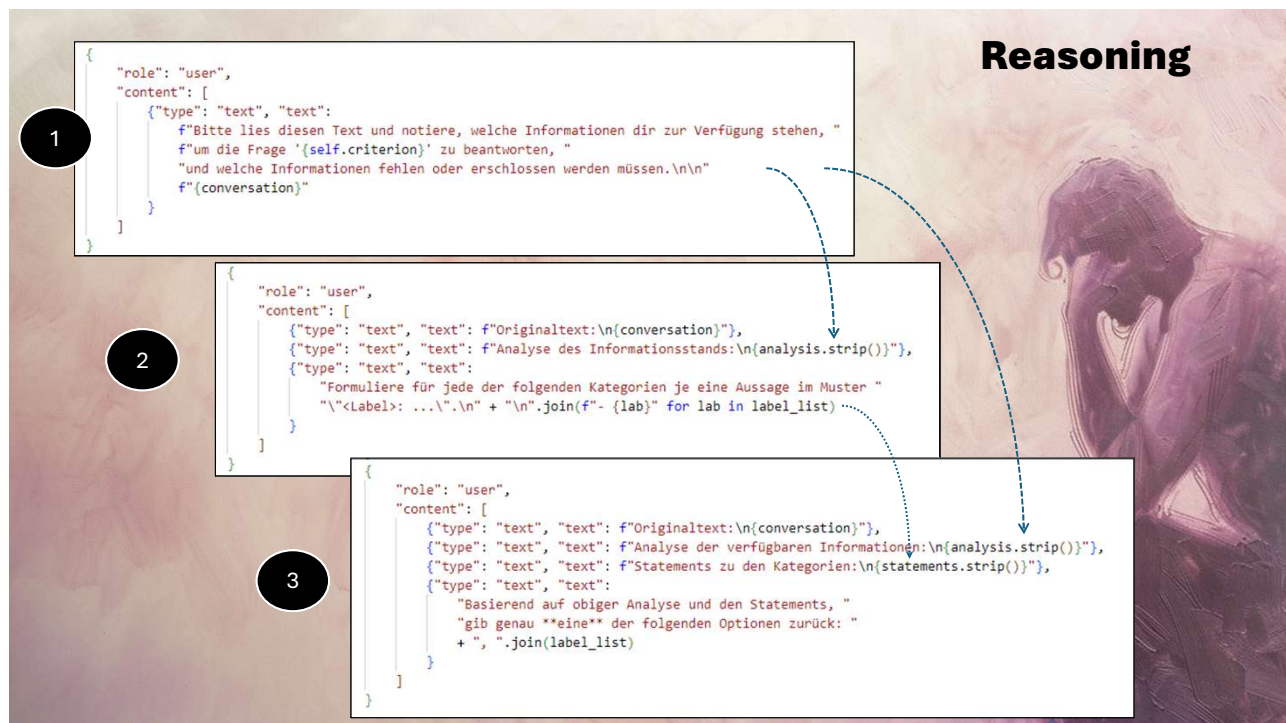
The entire conversation revolves around a specific man – Sam. Every reference to the wall, the photographs, the “cracks,” and the “things you try to conceal” are directly linked to his memory and the impact of his absence. Emilie’s meticulous measuring, her concern with “depth” and “layers,” and her lingering gaze at a photograph of him are all driven by the unresolved grief and attempts to grapple with his death. The dialogue explicitly mentions him (“a starting point. Like the first brushstroke. Do you think we should prime it first? A light grey, perhaps? To even out the... the shadows.”) and implicitly references him through the wall and the memories it holds. It’s a conversation entirely centered on Sam and the emotional baggage surrounding his death.

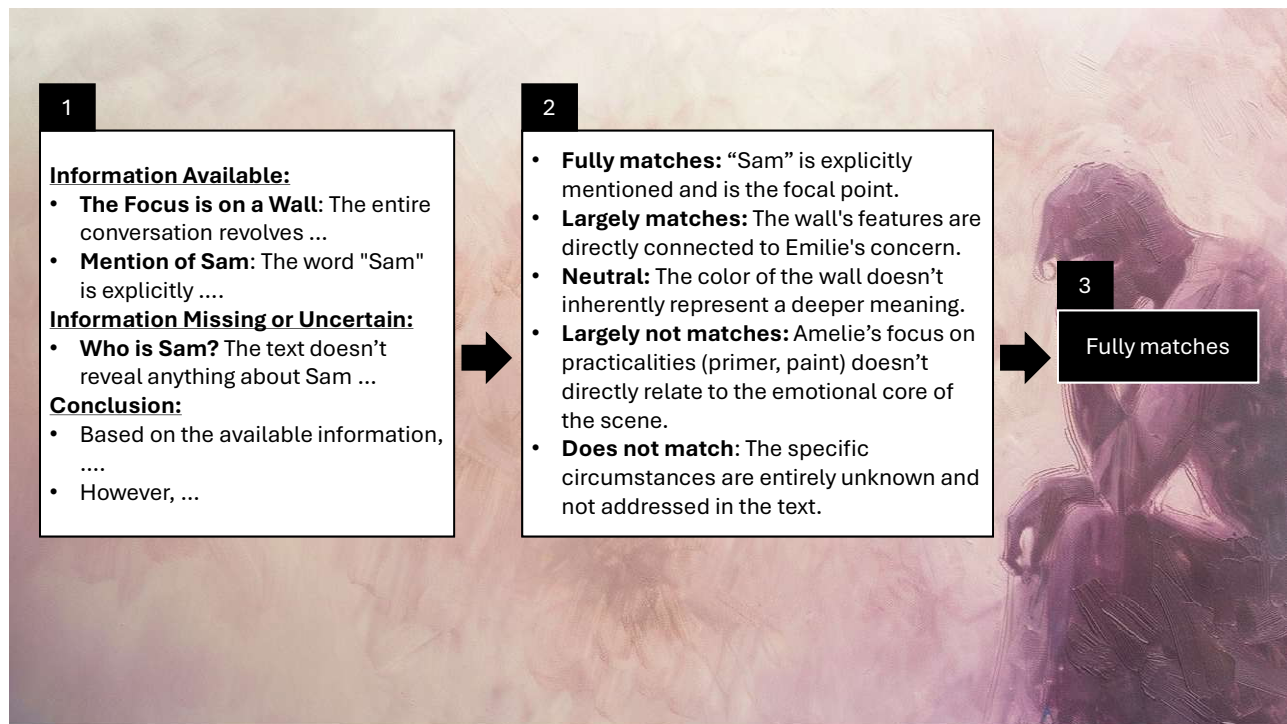
The primary subject of the conversation is something else than a man.

Fully matches

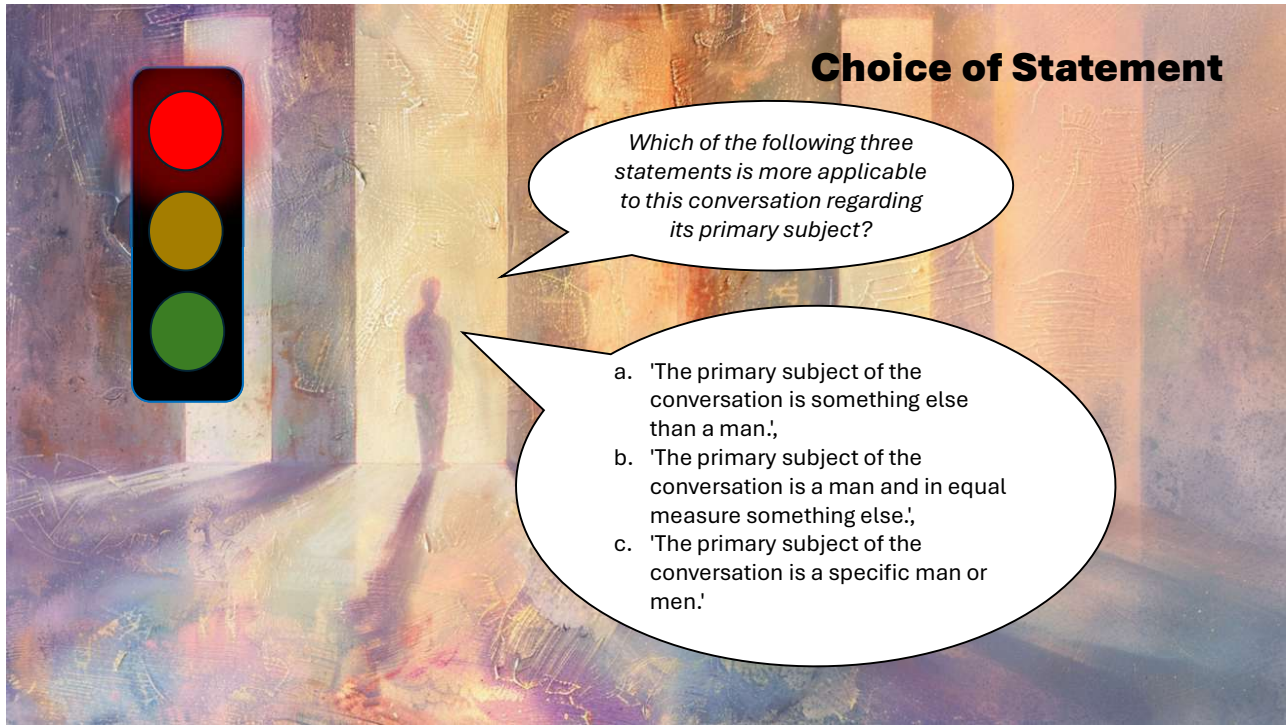
The conversation is overwhelmingly focused on the process of repairing a damaged wall. Every line of dialogue – the measurements, the discussion of primer and paint, the consideration of patching and texture, the reflection on “depth” and “cracks,” and the unspoken history hinted at by the photograph – centers entirely around the physical act of restoring the wall. The mention of “Sam” and the implication of concealed things are only relevant insofar as they relate to the wall’s history and the difficulty of confronting it. Therefore, the primary subject is demonstrably something other than a man.

- Statements cannot both be true?
- Explanations contradict each other!





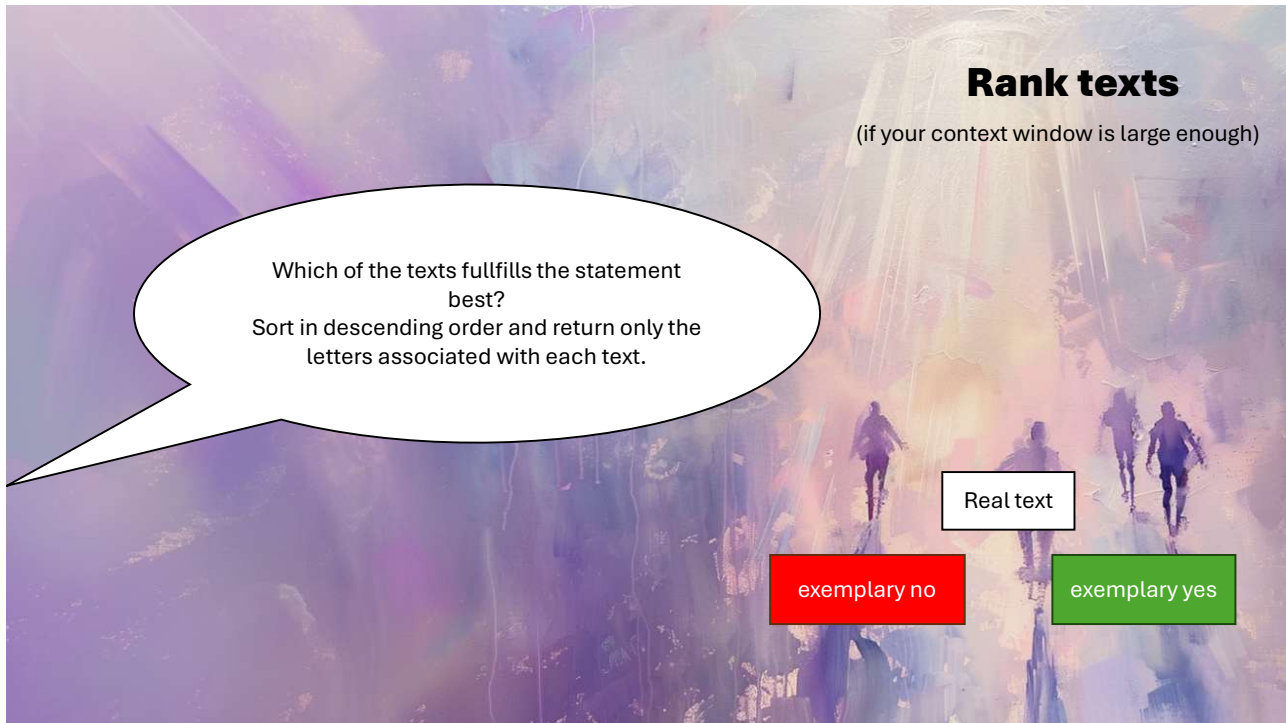
Bechdel-Wallace			
	Symbol	No Man	Just Man
This excerpt of a conversation between two women is mostly about a man or some men	★	★★	?
This excerpt of a conversation between two women is mostly about a topic that is not a man or some men.	★	★	★★



Choice of Statement

Which of the following three statements is more applicable to this conversation regarding its primary subject?

- a. 'The primary subject of the conversation is something else than a man.'
- b. 'The primary subject of the conversation is a man and in equal measure something else.'
- c. 'The primary subject of the conversation is a specific man or men.'



Rank texts

(if your context window is large enough)

Which of the texts fullfills the statement best?
Sort in descending order and return only the letters associated with each text.

Real text

exemplary no

exemplary yes



Training

Gemma3ForCausalLM

Text	Label
A kaleidoscope of neon signs and chattering crowds fills the street. Eleanor (20s), with a shock of bright pink hair pulled back in a messy ponytail, and Vivian (20s), sporting a stylishly messy blue bob, are strolling along, laughing. Rover, a fluffy golden retriever, trots happily beside them, tail wagging furiously ...	★ ★
Marcia: (With a dramatic sigh) ...and then Cousin Bartholomew, bless his cotton socks, tried to convince me he'd discovered a new species of beetle in the Cotswolds. A "beetle", Anna! As if that's the most pressing matter in the world. ANNA: (Dryly)...	★
Lucie: I am hungry Lara: So am I. Lucie: Let's eat. Lara: Pizza? Lucie: Well ... yes. But after that, I would like an ice cream. Lara: That is a plan.	★

Style	Intro
1940s Screw ball comedy	In a sun-dappled living room. There is a table with a vase of flowers and two glasses of martini.
Danish Dogme 95	In a stark, minimalist apartment, two friends engage in a heated discussion about life choices.

Generate a dialogue scene in the style of {style} based on the following scene description:
 {brief_desc}
 {full_desc}
 Stick to the scene description and the style, don't add additional personalities or plot.
 Dialogue:"

Text generation



Eleanor and her friend Vivian exchange warm words about Rover, Eleanor's devoted dog. But when Eleanor discovers a telltale lipstick stain on her husband's jacket, the light-hearted chatter about loyalty takes on a quietly ironic edge. With a brave smile and a tender pat for Rover, Eleanor keeps her secret—and her composure—intact.



Sandra and Lara are discussing their food plans. Both are hungry. Sandra is considering a Pizza. Lara would actually like something more healthy, but is settling on Pizza, because it is fast and she really is hungry. Sandra finds out, that they could also have ice cream. So they settle on a menu: First Pizza (with some alibi tomatoes on top) and then ice cream.



Sally and Diane are both very angry at Dave. When Sally starts insulting the absent Dave, Diane joins heartily. They keep adding up insults about Dave, until they both start to laugh.

Sunlight streams through sheer curtains, dappling the plush velvet furniture. A table sits center, holding a delicate vase of lilies and two impeccably chilled martinis. ELEANOR (impeccably dressed, a hint of forced brightness) and VIVIAN (equally stylish, a touch of mischievousness) are perched on a chaise lounge, deep in conversation. Rover, a stout, dignified bulldog, snores contentedly at Eleanor's feet.

VIVIAN: *(Taking a delicate sip of her martini)* Honestly, Eleanor, you've never met a dog quite like Rover. Such unwavering devotion. It's...remarkable. Like a furry, four-legged shadow.

ELEANOR: *(Smiling warmly)* Oh, he is. Absolutely. He's been with us through thick and thin, you know? A true companion. He follows me everywhere. Even to the grocery store.

She gently strokes Rover's head. He grunts in response.

VIVIAN: Good morrow, Eleanor! A sight most fair, To see thy Rover's grace, a joy to share! He doth possess a loyalty profound, A steadfast heart, upon most solid ground.

ELEANOR: *(Smiling, stroking Rover's head)* Indeed, dear Vivian, 'tis a truth well said. He's mine alone, a comfort, and my bread,

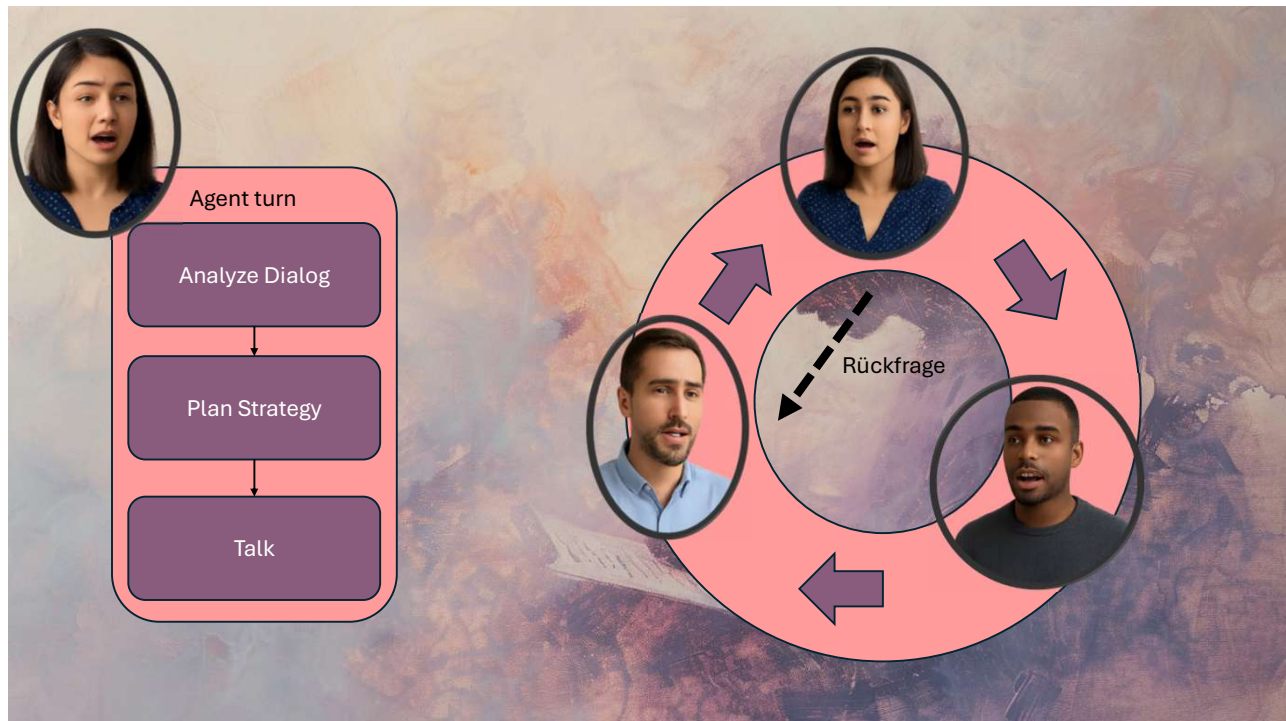
The room is grey. Bare walls. A single, worn armchair. A small, wooden table. ELEANOR sits with VIVIAN. A dog bed, Rover, is near the armchair.

VIVIAN: He's a good one, isn't he? That dog. Just... there.

ELEANOR: *(Smiling)* Yes. Always. Like a weight. A good weight.

(She strokes Rover, who sighs contentedly.)

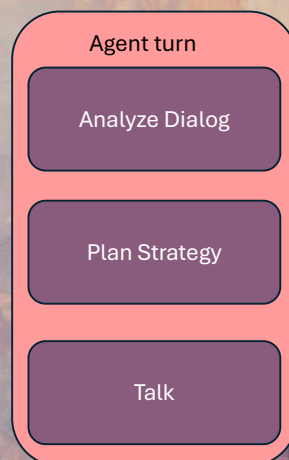




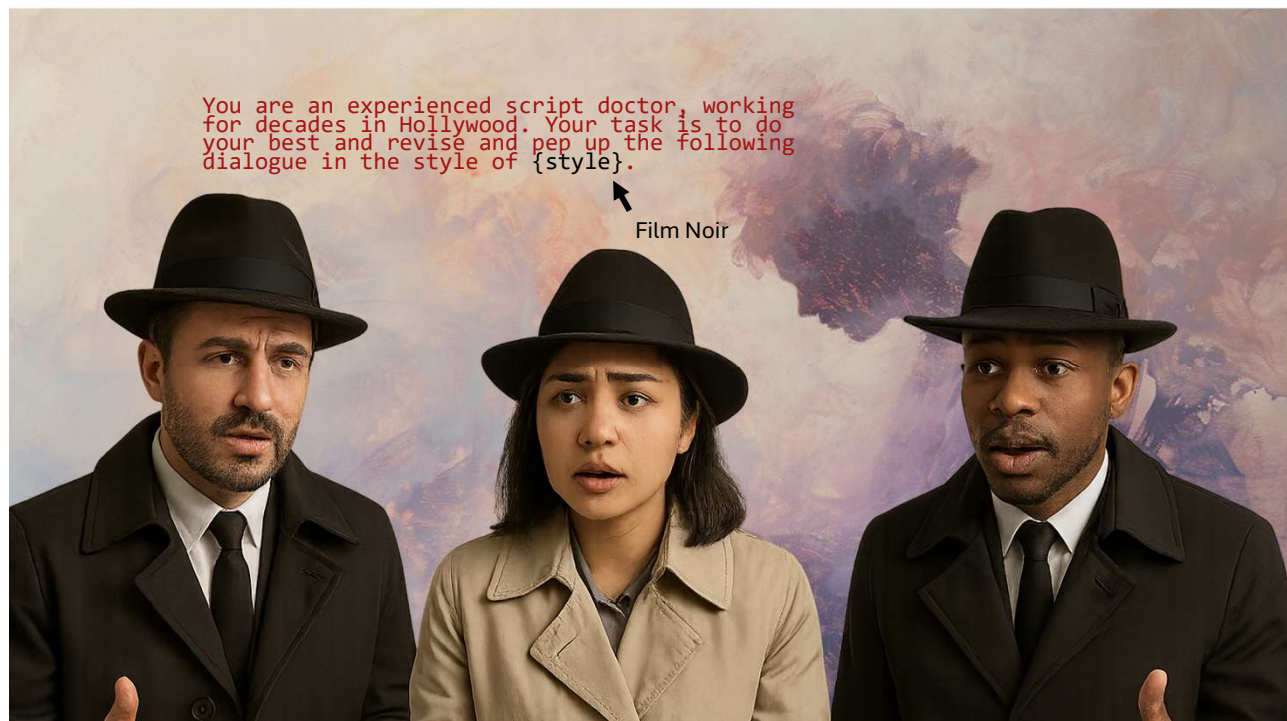
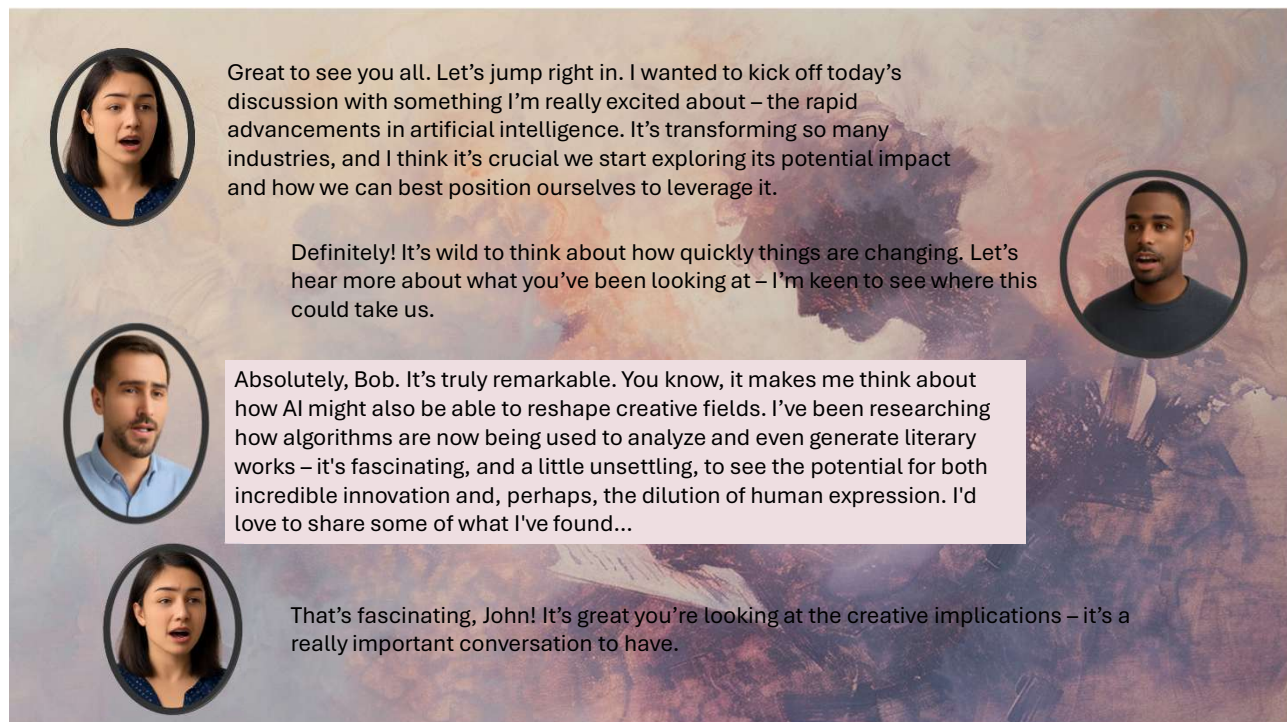
```
agent1 = Agent(
    name="Alice",
    topics=["technology", "artificial intelligence", "robotics"],
    role_desc="Alice ist technische Expertin mit einer Leidenschaft für Innovation.",
    special_actions={"summary": 0.1, "probe": 0.05}
)
```

Goals:

- Talk about all topics
- Don't be rude



- Greet
- Support last Talker
- Change Topic
- Continue Topic
- End Topic
- Start Topic
- Summarize previous Talk (Special)
- Address chosen Agent with question (Special)
- End Talk





I wanted to lay out something... significant. Artificial intelligence. It's not just a trend, it's a shift. A tectonic one. We need to understand how to use it, *control* it."

(Swallowing hard)
It's... unsettling, isn't it? Like looking into a mirror that shows you a future you don't want.



(Turning his gaze towards the window)
It's more than unsettling, Bob. It's a mimicry. They're feeding algorithms with *our* stories, *our* art... and then spitting back a pale imitation. I've been digging into literary analysis – these programs are already crafting prose. It's... efficient. And terrifying.

He pauses, a shadow crossing his face.
Like a forgery, but one that can convince you it's the original.



(Her voice dropping slightly) Precisely. We need to be ahead of the curve. A strategic advantage.



Good morrow to you all! 'Tis true, I muse,
That art itself may yield to cunning use.
These algorithms, weaving word and hue,
Do mimic minds, and strive to imitate you.
A potent grace, yet fraught with subtle dread,
To see our craft by cold mechanics led.
A shadowed beauty, born of logic's sway,
Where human touch may slowly fade away.
I have explored this labyrinthine quest,
And bring to light a truth, to be confessed.

