

Human-AI Collaboration

How can we think clearly ...
... about a technology whose consequences are still unfolding?

Alexandre de Spindler, desa@zhaw.ch
ZHAW Institut für Wirtschaftsinformatik

Chapters

1 Motivation

2 Approach

3 Foundations

4 Interacting with LLMS

5 Your Turn

Chapter 1: Motivation



Uncertainty: Can we see where this is going?



Orientation: Generative AI is just a piece of a larger development



Agency: We can learn to observe, hypothesize, measure, and adapt



Our current view
defines what we
notice ...

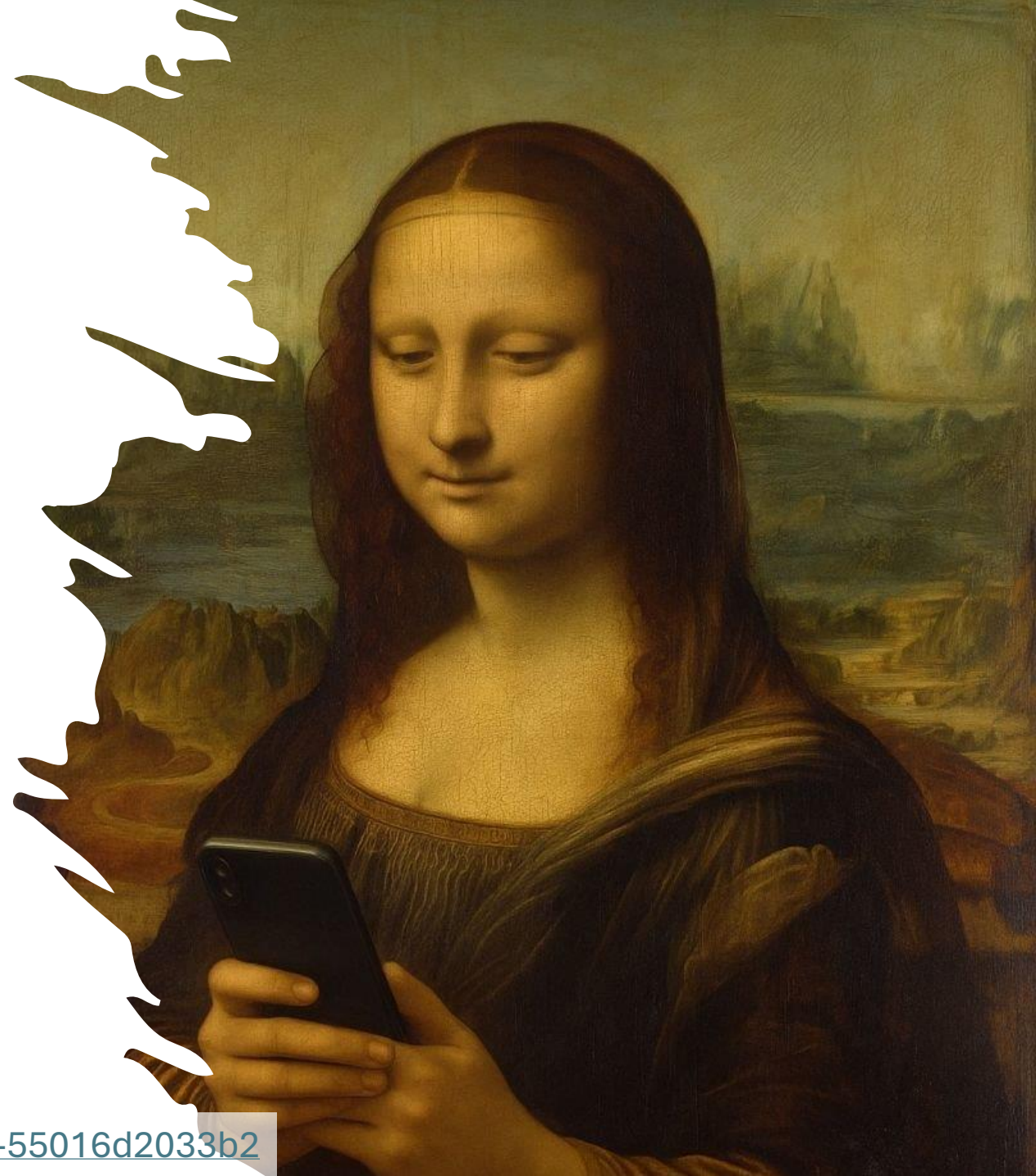
... and limits
what we can
imagine.



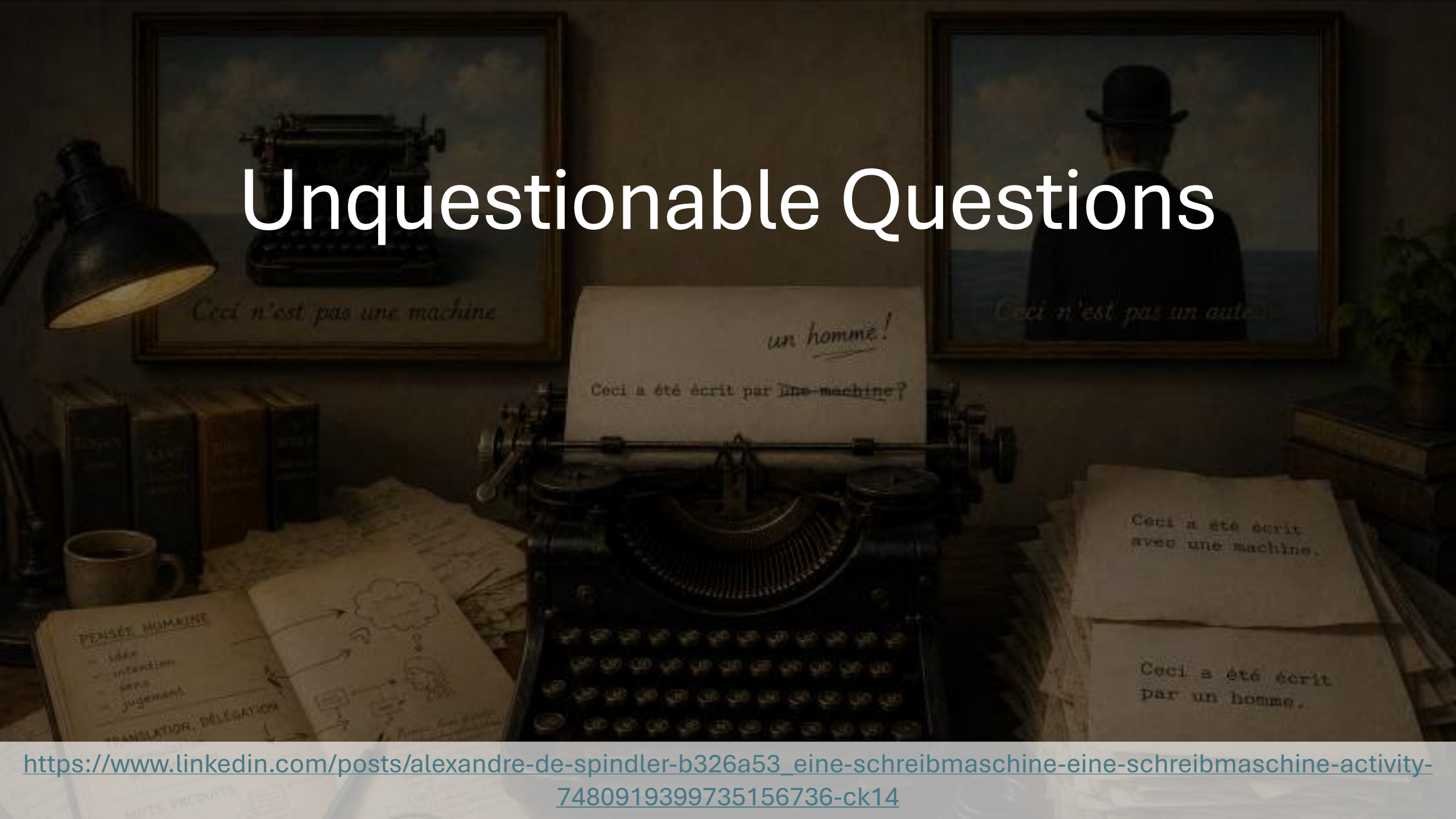
Social Media
Manager?

We tend to be
wrong about
predictions ...

... in so many ways!



Unquestionable Questions



Prompt-Assistenz der Geschäftsleitung

Ihre Aufgaben:

Sie bereiten KI-Sitzungen für wiederkehrende Geschäftsbedürfnisse vor.

Sie entwickeln Prompt-Schemata: also nicht nur einzelne Vorlagen, sondern wiederverwendbare Interaktionsmuster über mehrere Schritte, Rollen und Sitzungen hinweg.

Sie pflegen Kontextdokumente, damit relevante Informationen über Strategie, Produkte, Kunden, Prozesse und Entscheidungslogiken konsistent verfügbar bleiben.

Sie übersetzen Geschäftsbedürfnisse und -praktiken in geeignete Formen der Mensch-KI-Zusammenarbeit.

Sie achten darauf, wann KI nur ausführt, wann sie klärt, wann sie mitkonstruiert, wann sie kritisch reflektiert und wann mehrere KI-Interaktionen orchestriert werden müssen.

Sie beobachten, wie sich Modelle, Tools und Unternehmenskontext verändern, und passen die Arbeitsweise entsprechend an.

Ihr Profil:

Sie verstehen generative KI nicht nur als Werkzeug, Datenzugriff, und mathematisches Modell, sondern auch als Kollaborationssystem.

Sie können Aufgaben strukturieren, Kontext kuratieren, Interaktionen gestalten und Ergebnisse kritisch prüfen (lassen).

https://www.linkedin.com/posts/alexandre-de-spindler-b326a53_stellenanzeige-prompt-assistenz-der-gesch%C3%A4ftsleitung-activity-7457731261814710273-tv90



A milestone of
human
progress





Where is the seemingly accelerated development taking place?



Why would you be impressed about AI being part of a solution?
Why would that matter even more than the solution itself?

AI can't do that!

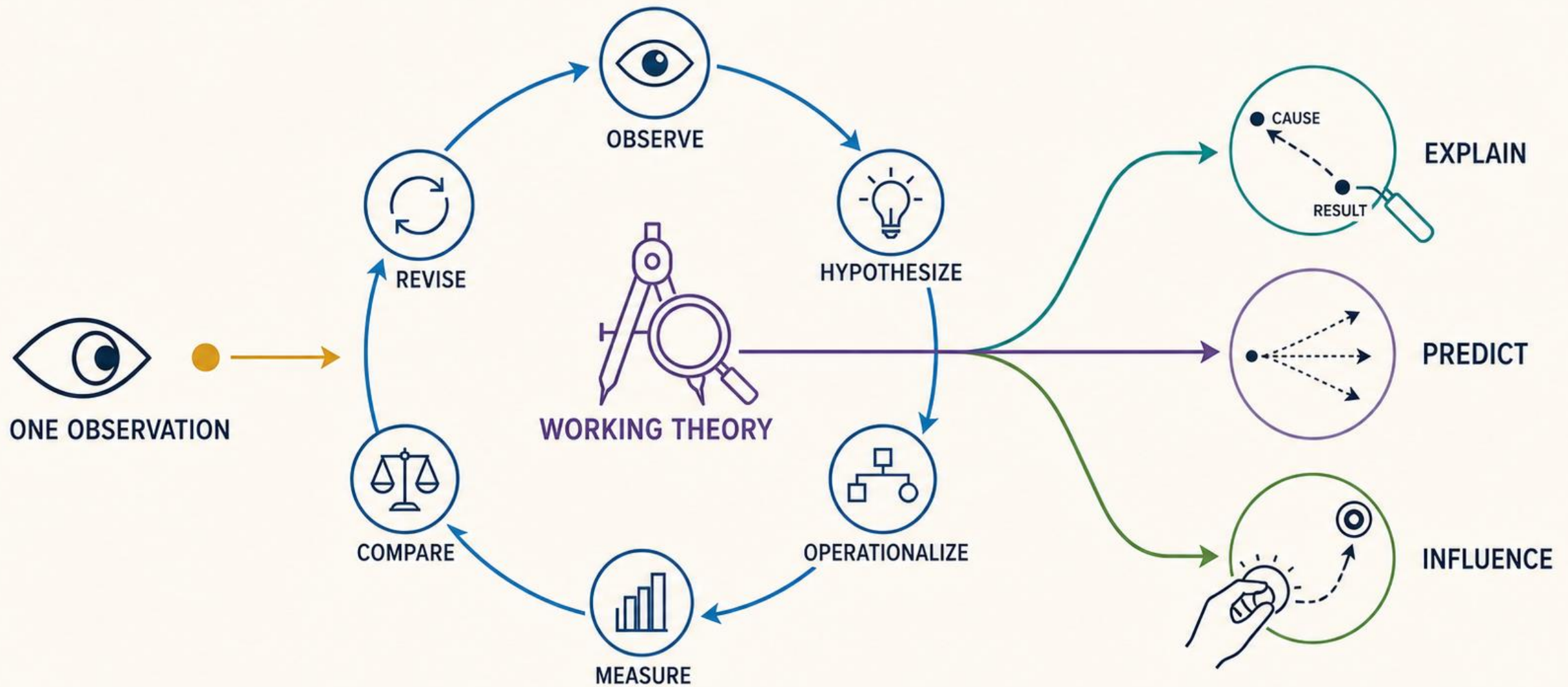
https://www.linkedin.com/posts/alexandre-de-spindler-b326a53_i-meiner-aktuellen-mediumcom-bubble-h%C3%A4ufen-activity-7452837788120502272-jdDJ



A photograph of a red pawn and a group of yellow pawns on a wooden surface. The red pawn is on the left, and the yellow pawns are on the right. The background is a blurred grey.

A Metaphor

Think of AI as a teammate



Boost your
effectiveness by
keeping your
working theory
up to date

https://www.linkedin.com/posts/alexandre-de-spindler-b326a53_ein-neues-huhn-ei-problem-in-der-bildung-activity-7430178719706546177-46NI

<https://medium.com/@despindler/can-generative-ai-make-us-smarter-4b81f8c2a94e>

Chapter 2: Approach



Collaboration levels: How we work with AI



Literacies: What we observe and control



Fundamentals: Why models behave this way



Challenges: Learn through controlled comparisons

Metaphors: useful views, claims to test

<https://www.youtube.com/watch?v=wv779vmyPVY>

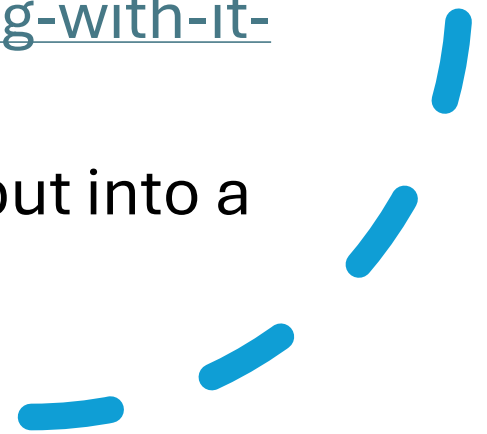
The limiting factor is no longer what AI can do, but how we choose to work with it.

<https://www.youtube.com/watch?v=loujaeBy8p0>

This isn't about clever prompts. It's about designing a thinking system around AI.

<https://medium.com/@despindler/stop-asking-ai-for-answers-start-working-with-it-cc829f828ef8>

All of the above, plus some more, put into a structure: collaboration levels



6 Collaboration Levels

0. Asking

The human asks for an answer with little explicit structure or control.

1. Instruction

The human specifies the goal, context, constraints, evidence expectations, and response format.

2. Alignment

Human and AI clarify assumptions, decisions, audience, and success criteria before generation.

3. Co-Construction

Human and AI shape the result together through purposeful phases, intermediate artefacts, and checkpoints.

4. Reflection

Construction, critique, revision, and verification are separated to identify weaknesses and improve the result.

5. Orchestration

The human designs control across multiple steps, tools, roles, permissions, resources, and AI interactions.



We can't
control what
we don't
understand.

Different explanations highlight different parts of model behaviour:

- 🚀 Say the right words, and the model will obey.
- 🚀 It's all just matrices and probabilities.
- 🚀 It's all in what they've been fed.

Each view explains part of what we observe. Use several perspectives to form and test predictions. None gives complete control.

5 Literacies



Interaction

How language, structure, context, examples, constraints, and feedback steer AI behaviour.



Mechanism

How computational mechanisms such as tokens, attention, context windows, and sampling shape AI behaviour and variability.



Data

How training data, sources, retrieval, and grounding shape what AI can produce and substantiate.



System

How workflows, tools, memory, permissions, handoffs, and checkpoints shape AI-supported work beyond a single interaction.



Human Context

How values, culture, responsibility, power, trust, and legitimacy determine whether human-AI collaboration is appropriate and successful.

Andelfingen

An aerial photograph of the town of Andelfingen in Switzerland. The town is nestled in a valley, surrounded by green hills. In the foreground, there are several houses with red-tiled roofs and a church with a tall, dark steeple. A railway line runs along the right side of the town. The background shows a wide river and distant hills under a clear sky.

Kann ein LLM verstehen, was Andelfingen mit Delfinen zu tun hat?

Was könnte das Wort «Keingang» bedeuten?

Kann ein LLM der Aufforderung "Wieviele Buchstaben hat Andelfingen" folgen?

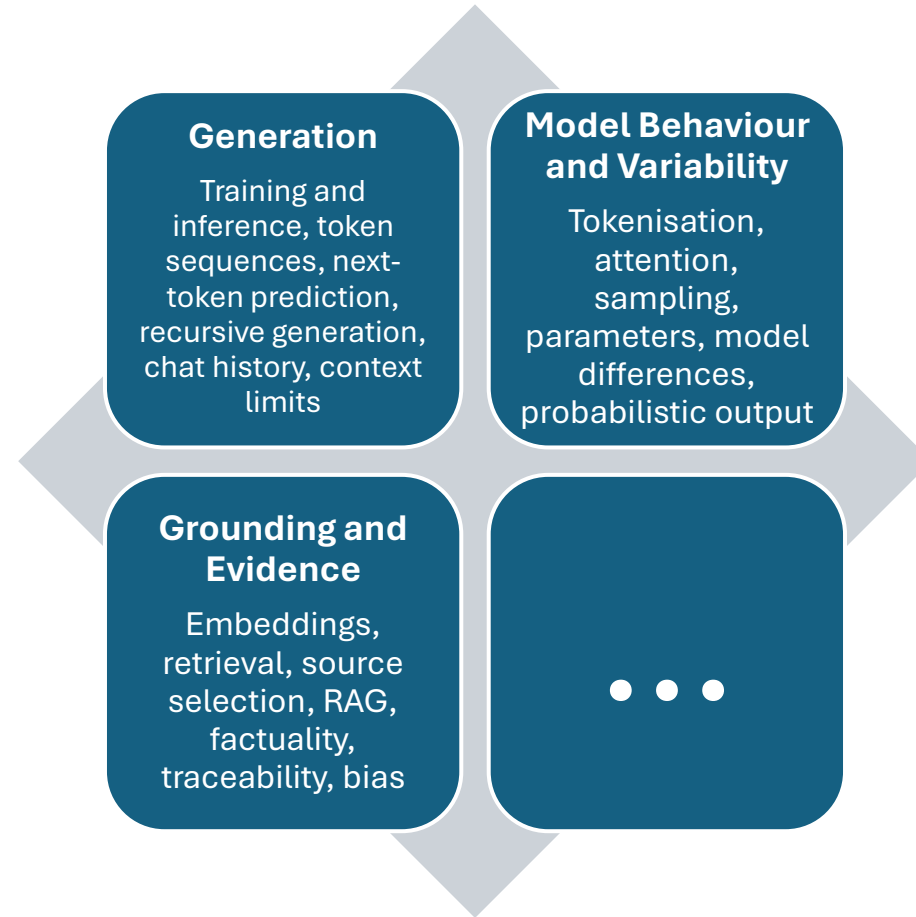
Welche Bedeutung hat ein mehrdeutiges Wort in diesem Satz?

Der Tokenizer zerlegt den Text.
Die Teilstücke hängen vom Modell ab.
Bekannte Textmuster können eine Deutung nahelegen.

Eine plausible Antwort belegt noch nicht, dass das Modell richtig gezählt hat.

Der Kontext hilft, die gemeinte Bedeutung einzugrenzen.

Fundamentals



Mastery



Challenges

- It develops and demonstrates a defined form of mastery.
- It requires effort, initiative, choices, adaptation, and sometimes persistence through failure.
- The goal is clear and learners can begin, but the route is not completely prescribed.
- It transfers the unit's learning into a relevant professional, disciplinary, or personal context.
- Learners predict an effect, manipulate or compare something, and observe what happens.
- Conclusions must be supported by what happened, not merely by opinion.
- Learners explain what happened, why, what remains uncertain, and what they would transfer to future work.
- 12h per Challenge, Pass/Fail

Chapter 3: Foundations



LLM Input & Prediction



From Token-Prediction to Individual Responses

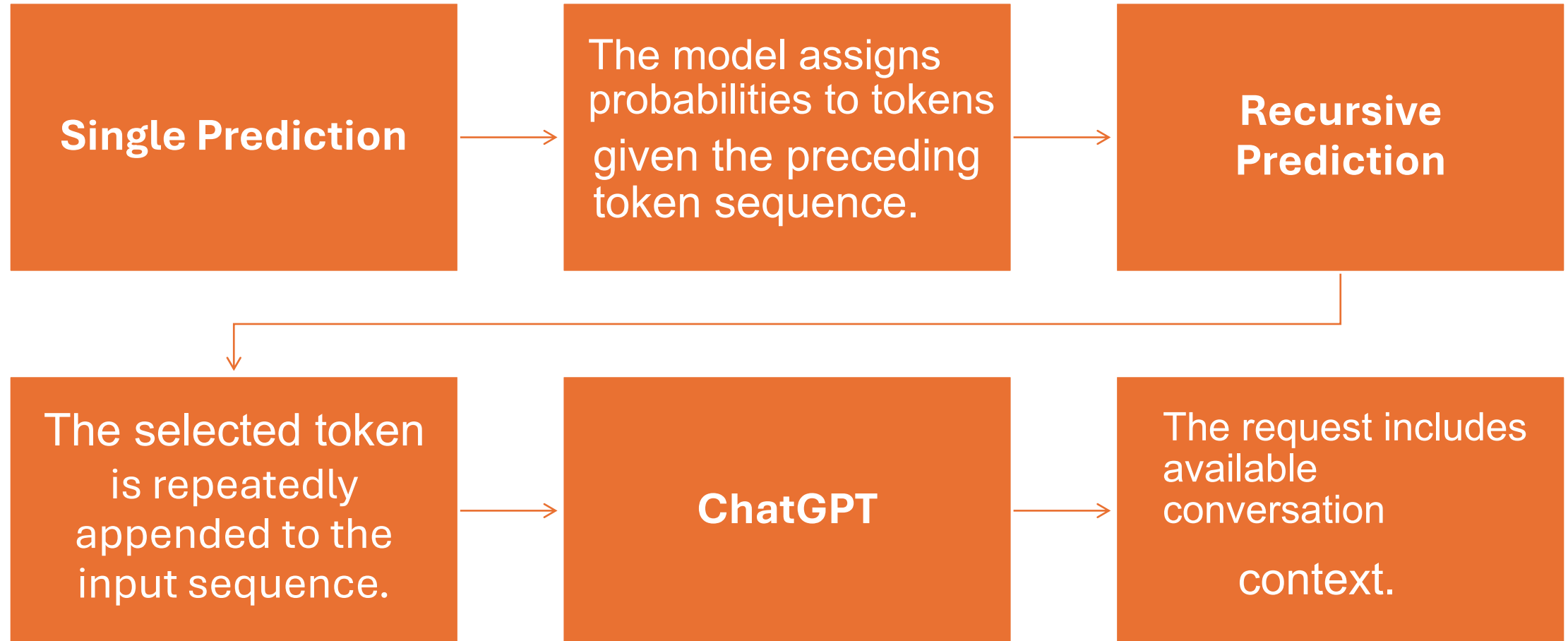


Chat History



Control: Training, Fine-Tuning, Prompt Engineering

Large Language Model (LLM) Fundamentals



1. Fundamental Capability (GPT)

Stimulus

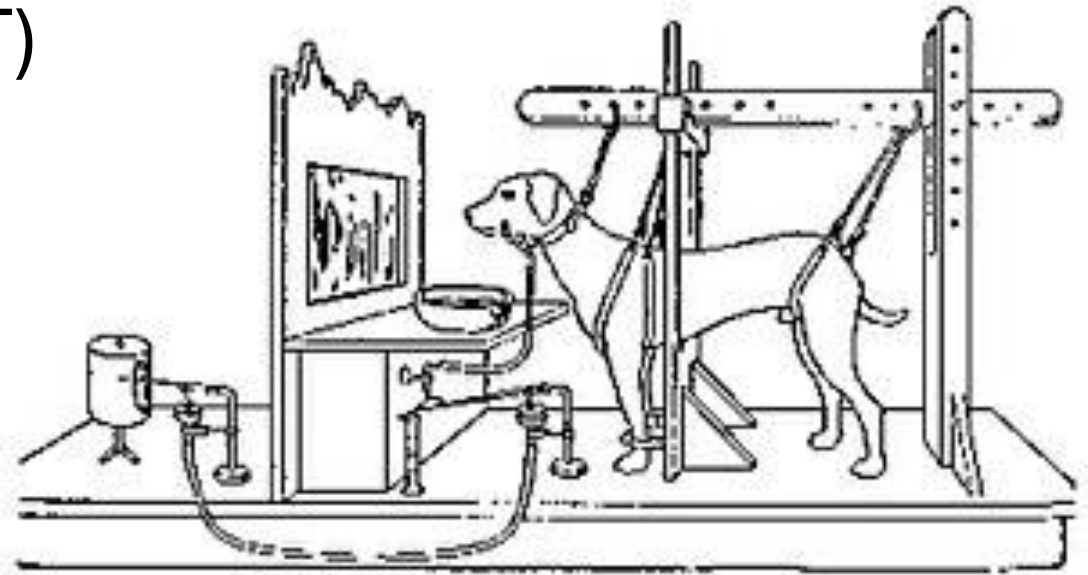
Sequence of text pieces (tokens)

["I", "could", "probably"]

Response

One Token

"beat"



2. Recursive Prediction (GPT Completion)

Append the selected token and predict again

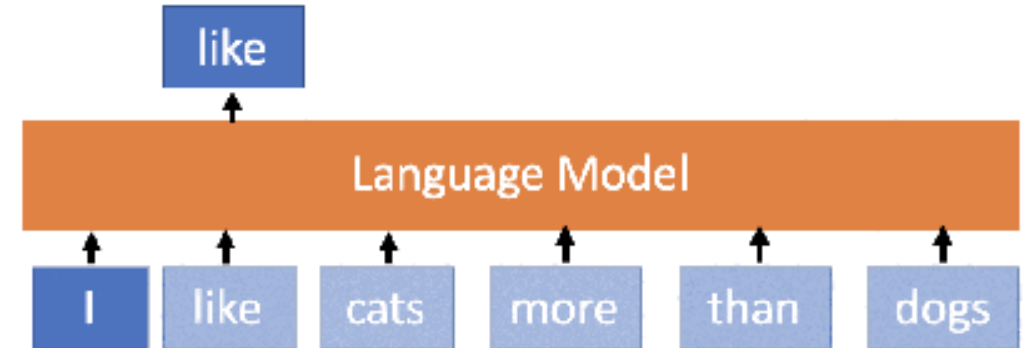
```
["I", "could", "probably"] > "beat"  
["I", "could", "probably", "beat"] > "you"  
[..., "probably", "beat", "you"] > "in"  
[..., "beat", "you", "in"] > "chess"
```

...

Until ...?

The output token limit is reached

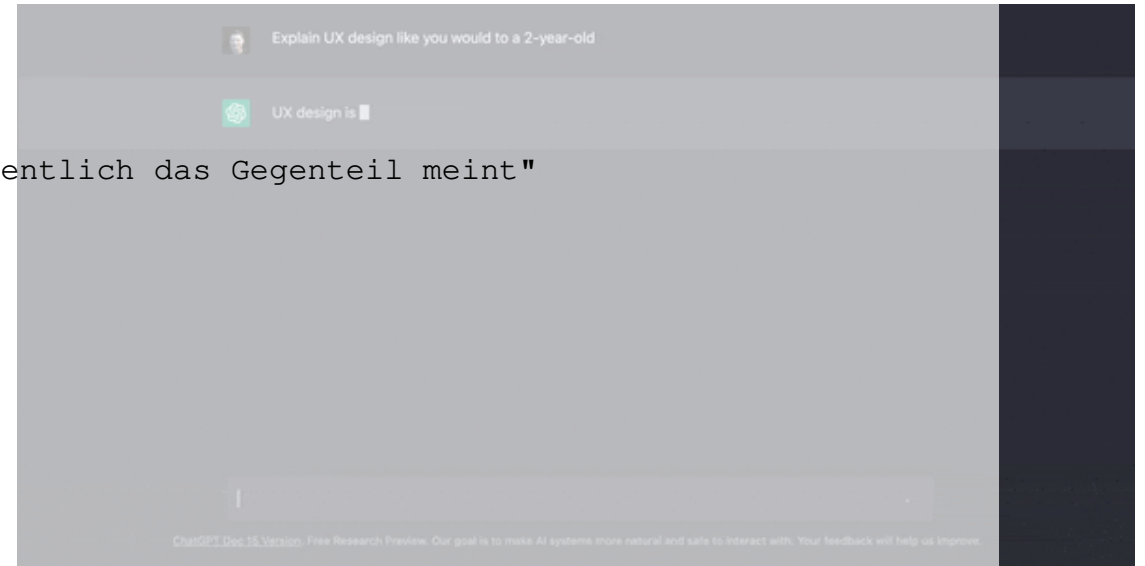
An end-of-sequence token is selected



3. Chat History (ChatGPT)

A chat request includes selected conversation context. Example:

```
[
  {
    "role": "system",
    "content": "Du beantwortest meine Fragen als wäre ich fünf Jahre alt"
  },
  {
    "role": "user",
    "content": "Was ist Ironie?"
  },
  {
    "role": "assistant",
    "content": "Ironie ist, wenn man etwas sagt, aber eigentlich das Gegenteil meint"
  },
  {
    "role": "user",
    "content": "Sag mal nein"
  },
  {
    "role": "assistant",
    "content": "Nein!"
  },
  {
    "role": "user",
    "content": "War das jetzt ironisch?"
  }
]
```



<https://www.thefountaininstitute.com/blog/chat-gpt-ux-design>

How to control model behaviour?

Pre-Training

Learn patterns from training data

Requires data and substantial computation.

Fine-tuning (a form of post-training)

Adapt a trained model with selected examples

Changes learned behaviour.

Still requires evaluation on the intended task.

At run time

Provide instructions, context and tools.

Test their effect in the actual workflow.

**Training and run-time inputs are different ways
to influence behaviour. Neither guarantees correctness.**

How to control model behaviour? ...

Prompt Engineering

Design the conversation and its checkpoints.

Set instructions and provide relevant context.

Example Structures

1. Role, Task, Format
2. Context, Task, Format
3. Role, Action, Steps, Context, Example, Format
4. ...

But ...

Long prompts can hide conflicting requirements.
Separate stages when this makes checks easier.

Prompt Engineering

Zero-/Few-Shot Learning

Providing Context

Chain-of-Thoughts Prompting

Role-Prompting & Multiple Persona

Program Simulation

Response Format

Iterative/ Expansive Prompting

Meta Prompting

<https://academy.openai.com/public/courses/ai-foundations-juzjs>
<https://arxiv.org/abs/2311.05661>



If there is only one
thing you learn, let
it be this

Use an LLM like this



Fragen beantworten

Zum Beispiel zu einem Begriff oder Ereignis. Sachangaben anhand von Quellen prüfen.



Inhalte entwerfen

Texte, Präsentationen, Interviewfragen, Anleitungen oder Ideen



Inhalte verdichten

Zusammenfassen, Themen oder genannte Personen und Organisationen herausarbeiten



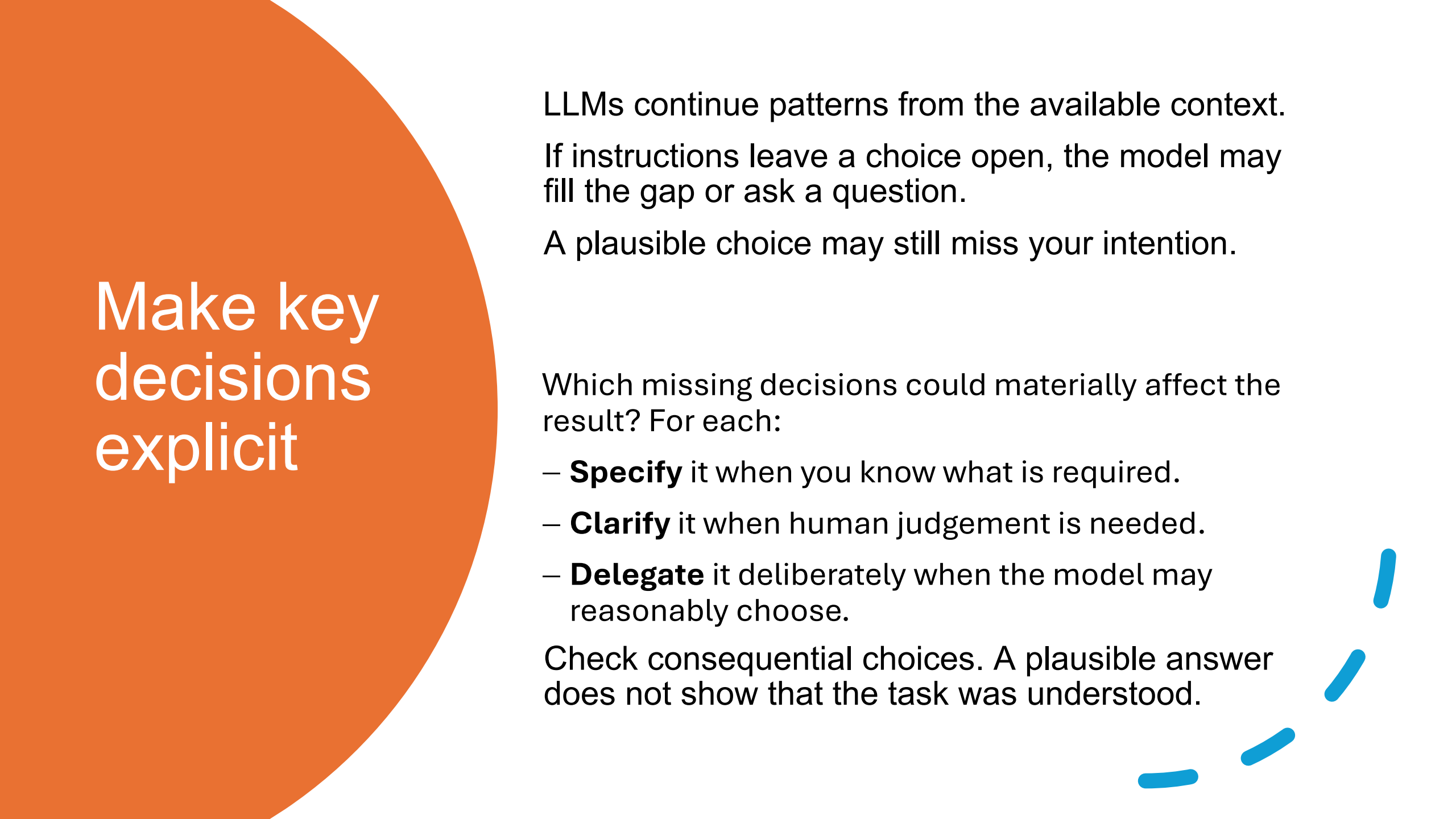
Inhalte umarbeiten

Erklären, übersetzen, umformulieren oder für eine andere Zielgruppe aufbereiten



Mit strukturierten Formaten arbeiten

Zum Beispiel Tabellen, Code, XML oder JSON erzeugen und prüfen



Make key decisions explicit

LLMs continue patterns from the available context.

If instructions leave a choice open, the model may fill the gap or ask a question.

A plausible choice may still miss your intention.

Which missing decisions could materially affect the result? For each:

- **Specify** it when you know what is required.
- **Clarify** it when human judgement is needed.
- **Delegate** it deliberately when the model may reasonably choose.

Check consequential choices. A plausible answer does not show that the task was understood.

Chapter 4: Interacting with LLMs



Let AI find out what needs to be known (Specification)



Let AI suggest an approach (Planning)



Let AI help you improve (Feedback)

Work with AI and check what changes



Try collaboration where it serves the task
Divide the work into stages. Compare the result and effort with a direct request using the same criteria.



Use AI suggestions to improve the interaction
Test its advice. The model is not an authority on its own reliability.

<https://medium.com/@despindler/stop-asking-ai-for-answers-start-working-with-it-cc829f828ef8>

Level 0: Asking for Answers

You give a short instruction. The system produces an answer.

```
Draft an outline for a digital-transformation initiative proposal.
```

Level 1: Instruction

Move from vague prompts to structured instructions.

```
Create an executive summary of the following analysis.
```

```
<analysis>
```

```
...
```

```
</analysis>
```

```
Your response should:
```

- focus on ROI
- highlight risks
- be limited to one page

Level 2: Alignment

First ensure we agree on the problem.

```
I need an executive summary based on the following analysis.
```

```
<analysis>
```

```
...
```

```
</analysis>
```

```
Before writing it, ask me questions to clarify:
```

- the audience
- strategic priorities
- decision context

Level 3: Co-Construction

Instead of asking for an answer, we co-design a process.

Act as a strategy consultant.

First, ask me 3 questions to understand:

- business objectives
- constraints
- stakeholders

Then design a step-by-step approach to create an executive recommendation, including analysis, prioritization, and communication.

Level 4: Reflection

We stop asking only "Is this a good result?" and start asking "How did we arrive here, and what might be wrong?"

Based on the recommendation:

- What are the biggest risks or blind spots?
- Where could this fail in practice?

Level 5: Orchestration

Coordinate stages with defined inputs, outputs and checks.
The sequence below is a starting point for a workflow.

```
Step 1: Analyse the following data.
```

```
<data>  
...  
</data>
```

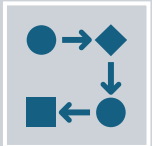
```
Step 2: Identify key insights and risks.
```

```
Step 3: Translate these into an executive recommendation.
```

Chapter 5: Your Turn



Task: Create an Artefact



Procedure: One-Shot vs. 5-Phase Approach



Evidence: How could you compare the two resulting artefacts?

In-Class Challenge

Bereiten Sie eine fünfminütige Präsentation zu einem zugelosten Thema vor.

Ihre Präsentation soll den anderen Kursteilnehmenden das Thema verständlich, faktisch fundiert und mit einer klaren Kernbotschaft vermitteln.

- **Zielgruppe:** die anderen Kursteilnehmenden
- **Dauer:** 5 Minuten
- **Umfang:** ungefähr 3–5 Folien
- **Artefakt:** Folien mit vollständigen Sprechhinweisen oder ausformuliertem Sprechtext

Ziel: Nach der Präsentation soll die Zielgruppe das Thema verstanden haben und die wichtigste Aussage wiedergeben können.

5-Phasen Interaktion

- 1. Kontext strukturieren**
Thema, Zielgruppe, Wirkung, Vorgaben sowie Fakten- und Zielgruppengrundlage bereitstellen.
- 2. Lücken klären**
Offene Fragen und notwendige Entscheidungen identifizieren und selbst beantworten.
- 3. Vorgehen vereinbaren**
Aufbau, Kernbotschaft, Folienfolge und Zeitverteilung prüfen und freigeben.
- 4. Vorgehen ausführen**
Folien und Sprechtext auf Grundlage des vereinbarten Vorgehens erstellen.
- 5. Prüfen und verbessern**
Fakten, Verständlichkeit, Zielgruppenpassung und beabsichtigte Wirkung prüfen und Schwächen korrigieren.

Was nützt's?

1. Vergleich vorbereiten
Dieselbe Aufgabe, Zielgruppe und beabsichtigte Wirkung.
Möglichst dasselbe Modell verwenden.
2. Zwei Fassungen erstellen
A: Direktanfrage ohne erarbeitete Zusatzgrundlage.
B: Fünf Phasen mit geprüften Grundlagen.
3. Mit vier Kriterien vergleichen
Richtigkeit: geprüfte Fakten.
Zielgruppenpassung: begründetes Profil.
Wirkung: vorher festgelegter Indikator.
Nachvollziehbarkeit: Aussagen auf Grundlagen zurückführen.
4. Schluss ziehen
Unterschiede und Unsicherheiten festhalten.
Eine begründete Verbesserung ableiten.