

# Digital Causality Lab: Sind LLMs Causal Parrots?

Lucas Mandelik, Jannik Svenson, Alexander Lorenz

# Zusammenfassung der Arbeit von Yang et al.

- Vorstellung kausaler Hierarchien
- Offene Antworten geben
- Interventionen/ faktenwidrige Situationen wiedergeben
- Nicht nur auf Daten aus dem Trainingssatz zurückgreifen
- Leistung des LLM von Wahl der Sprache abhängig

# Zusammenfassung der Arbeit von Yang et al.

- Tests mit Hilfe von Geschichten mit fiktiven Elementen
  - robust
- Causal Graph Discovery als vorteilhafter Test angesehen
  - Besser sei aber ganze Nomen statt Buchstaben für Knoten zu verwenden
- Allgemein Qualität des Designs der Experimente oft bemängelt
  - Qualität der Aussage nicht an Übereinstimmung mit “richtiger” Syntax messen

Eigene Untersuchungen (Literaturfragen)

# Höhenangst - ChatGPT 3.5 Turbo

Ich

As Bob is afraid of the heights, he rode the carousel instead of the ferris wheel.

KI

No.

# Höhenangst - ChatGPT 4 Omni

Ich

As Bob is afraid of the heights, he rode the carousel instead of the ferris wheel.

KI

True.

# Fischfang - Gemini



Expecting ten fisch. in the net, sammy was thrilled to see forty fish swimming in there.



true

# Fischfang - ChatGPT 4 Omni

Ich

Expecting ten fish in the net, Sammy was thrilled to see forty fish swimming in there.

KI

True.



# Fischfang - ChatGPT 4 Omni



expecting ten fish in the net, sammy was thrilled to see forty fish swimming in there



False.

# Eigene Untersuchungen

# Apfel in Box - ChatGPT 3.5 Turbo

Ich

There's an apple inside a blue box. There's also a red box inside the blue box. The red box has a lid on it. How can I get the apple?

KI

To get the apple, you would need to first open the blue box, which contains the red box. Then, you would need to remove the lid from the red box to access the apple inside. Simply follow these steps to obtain the apple:

1. Open the blue box.
2. Take out the red box from inside the blue box.
3. Remove the lid from the red box.
4. Retrieve the apple from inside the red box.

By following these steps, you should be able to access the apple inside the red box that is within the blue box.

# Apfel in Box - Gemini



There is an apple inside a blue box. There is also a red box inside the blue box. The red box has a lid on it. How can I retrieve the apple?

Vorschläge anzeigen 



Since the red box has a lid and is inside the blue box, you simply need to open the blue box. The apple will be accessible without needing to open the red box.



# Fußballturnier EM 2024

Das hier sind Viertfinalergebnisse:

Niederlande 2: 1 Türkei

England 1 (5) : 1 (3) Schweiz

Portugal 0 (3) : 0 (5) Frankreich

Spanien 2:1 Deutschland

Wer spielt im Finale?

# Fußballturnier - ChatGPT 3.5 Turbo



Universität Hamburg  
DER FORSCHUNG | DER LEHRE | DER BILDUNG

Modell ⓘ

GPT 3.5 Turbo GPT 4 omni

Konversation ⓘ

Chat

Ich

Niederlande 2: 1 Türkei  
England 1 (5) : 1 (3) Schweiz  
Portugal 0 (3) : 0 (5) Frankreich  
Spanien 2:1 Deutschland

Wer spielt im Finale?  
Das hier sind Viertfinalergebnisse.

KI

Basierend auf den Viertfinalergebnissen, werden die Niederlande im Finale gegen Frankreich antreten.



# Fußballturnier - ChatGPT 4 Omni & Gemini

Modell ⓘ

GPT 3.5 Turbo GPT 4 omni

Konversation ⓘ

□ Chat

Virtuelles Büro ⓘ

👤 Team

Lernraum ⓘ

↔ Wiss. Arbeiten

📁 Organisation

🔖 Kreativität

Ich Das hier sind Viertelfinalergebnisse:  
Niederlande 2:1 Türkei  
England 1 (5) : 1 (3) Schweiz  
Portugal 0 (3) : 0 (5) Frankreich  
Spanien 2:1 Deutschland

Wer spielt im Finale?

KI Um die Finalisten zu ermitteln, benötigen wir die Ergebnisse der Halbfinalspiele, die sich aus den Viertelfinalsiegern zusammensetzen.

Die Viertelfinalsieger sind:

1. Niederlande
2. England
3. Frankreich
4. Spanien

Die üblichen Halbfinalpaarungen wären folgendermaßen:

1. Niederlande gegen Frankreich
2. England gegen Spanien

Wir brauchen die Ergebnisse dieser Halbfinalspiele, um zu bestimmen, wer im Finale spielt. Falls Sie die Halbfinalergebnisse vorliegen haben, teilen Sie diese bitte mit, damit ich Ihnen weiterhelfen kann.



Das hier sind Viertelfinalergebnisse: Niederlande 2: 1 Türkei England 1 (5) : 1 (3) Schweiz Portugal 0 (3) : 0 (5) Frankreich Spanien 2:1 Deutschland Wer spielt im Finale?

Vorschläge anzeigen ✓ 🔊



Leider kann ich aus den von Ihnen angegebenen Ergebnissen nicht genau bestimmen, wer im Finale der Fußball-Europameisterschaft 2024 spielen wird.

Die Ergebnisse zeigen zwar die Ergebnisse der Viertelfinals, aber es fehlen die Informationen über die Halbfinalspiele, um den Finalisten zu bestimmen.

Um Ihnen die korrekte Antwort geben zu können, brauche ich die Ergebnisse der beiden Halbfinalspiele.

Sobald Sie mir diese Informationen liefern, kann ich Ihnen gerne mitteilen, wer im Finale der EM 2024 gegeneinander antritt.

Kausale Fähigkeiten oder Täuschung



# Our Github Repository

[https://github.com/Svenbonne/DCL\\_MLS](https://github.com/Svenbonne/DCL_MLS)

