

Named Entity Recognition mit LLMs

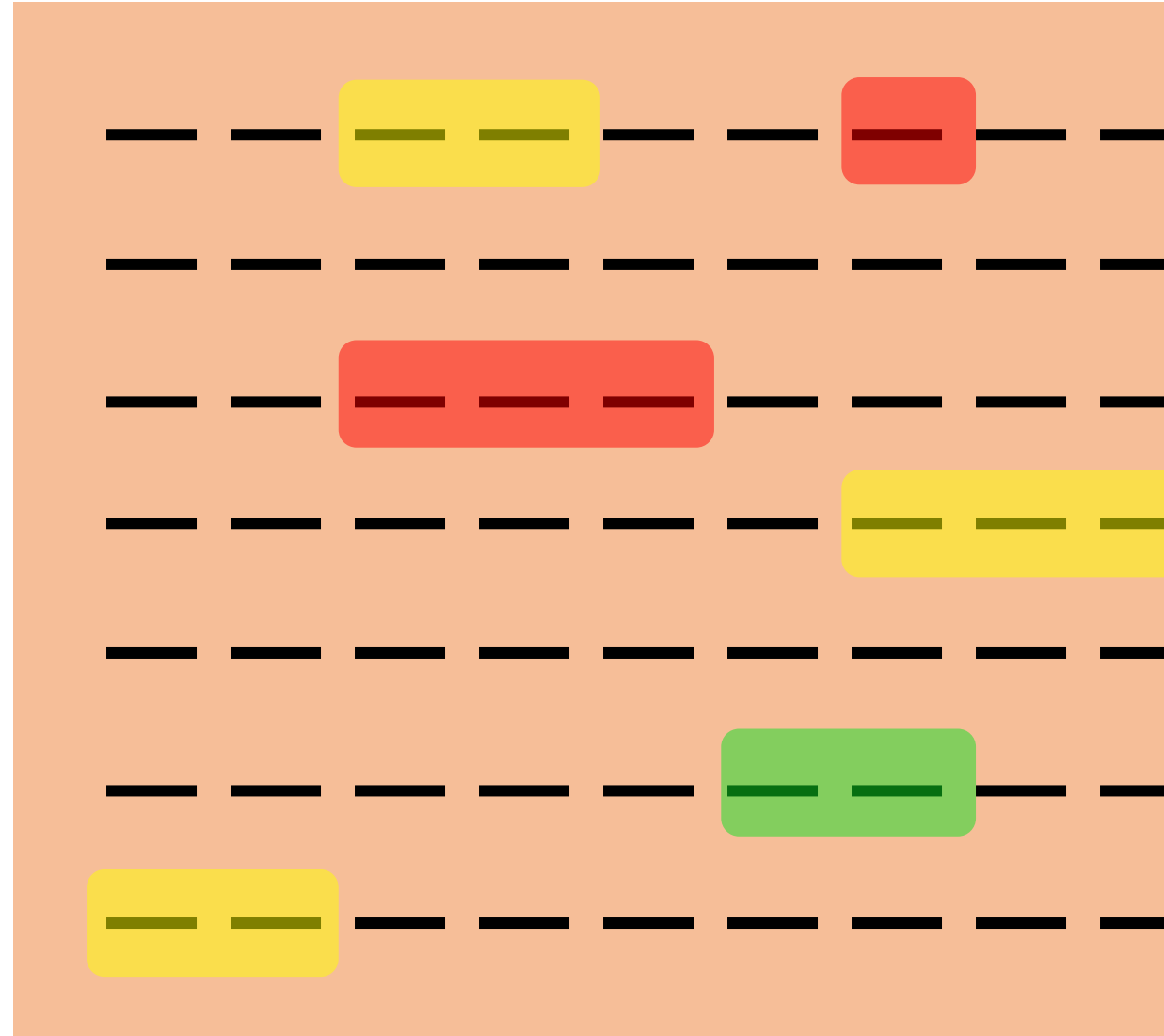
Nina C. Rastinger

Webinar der DHd-AG „Angewandte

Generative KI“, 25.09.2024

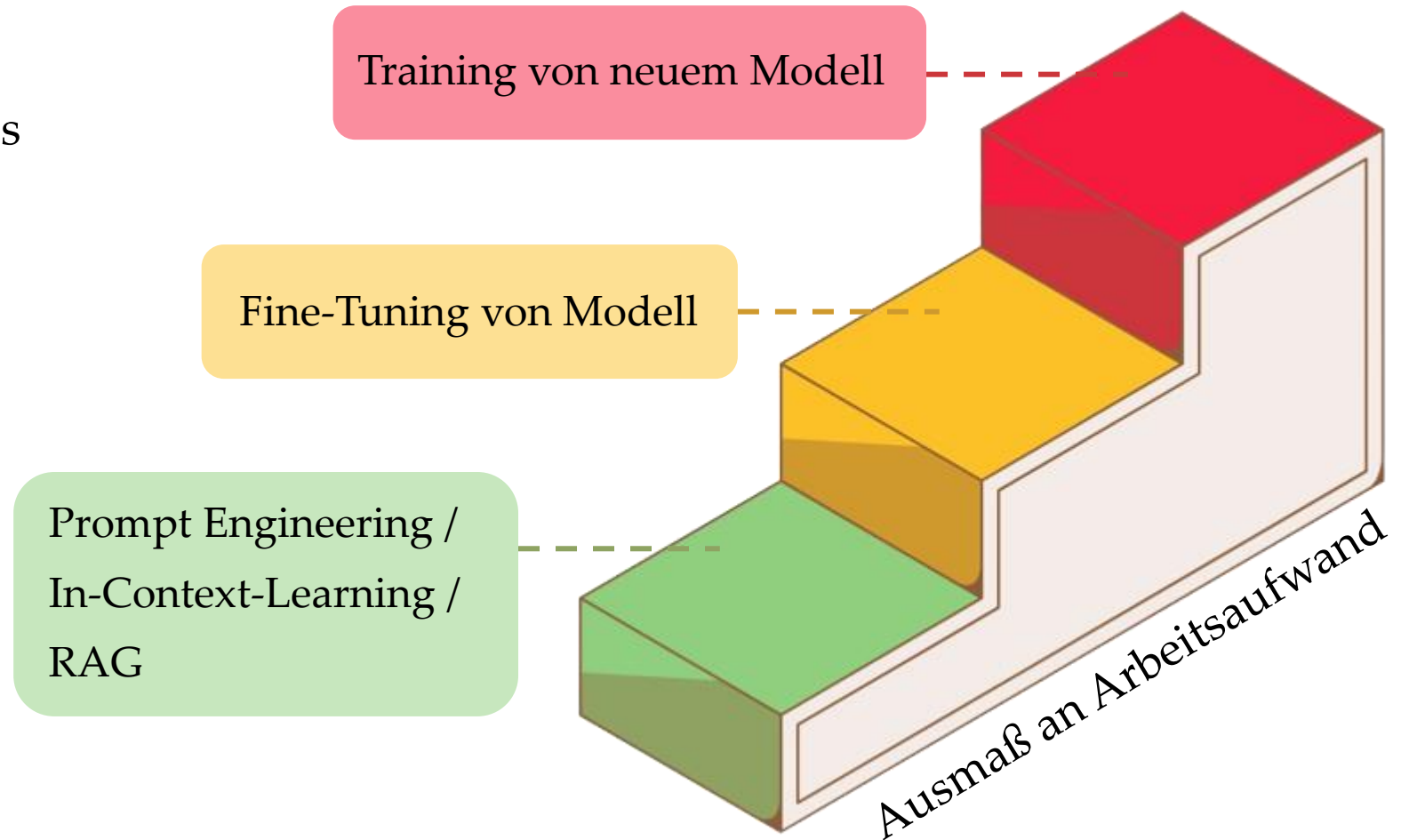
Hintergrund

- Eigene Anwendungsfälle und Experimente mit Listen in historischen Zeitungen und Reisehandbüchern
- Review von 18 Studien zu NER & GPT (aus DH- und Nicht-DH-Bereich)
- AG „Angewandte Generative KI“ > Versuch der Ableitung praktischer Hinweise
- Weg Richtung „Best Practices“: Suche nach entscheidenden „Erfolgsfaktoren“ für NER mit LLMs



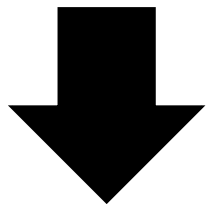
Hintergrund

- Verschiedene Stufen der Anpassung von LLM an eigenes Datenmaterial
- Idee hier: möglichst niederschwelliger Zugang – Optimierung vor Wechsel zu nächsthöherer Stufe

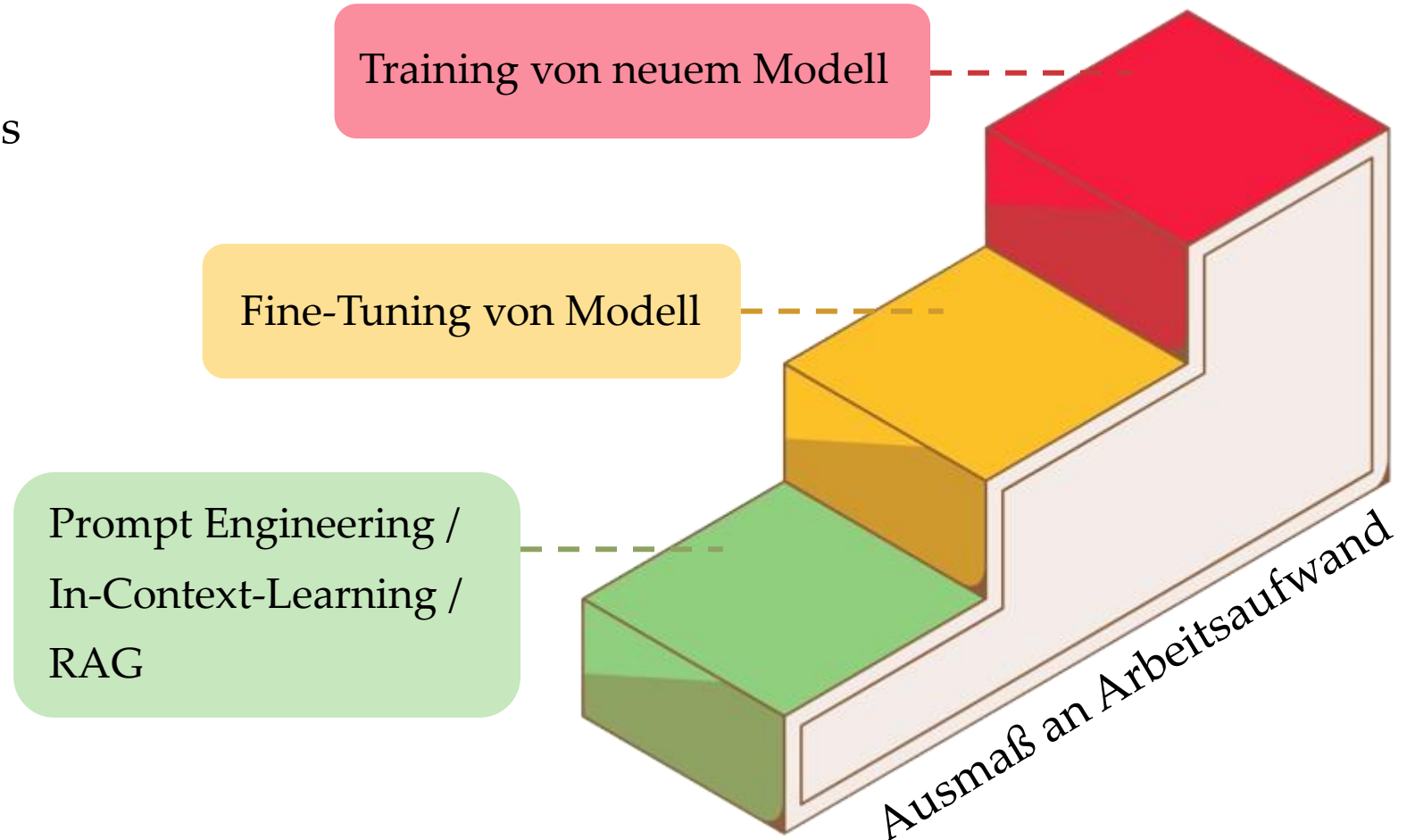


Hintergrund

- Verschiedene Stufen der Anpassung von LLM an eigenes Datenmaterial
- Idee hier: möglichst niederschwelliger Zugang – Optimierung vor Wechsel zu nächsthöherer Stufe



bereits auf unterster Stufe hohes Erfolgspotenzial durch Vielzahl an Eingriffsmöglichkeiten

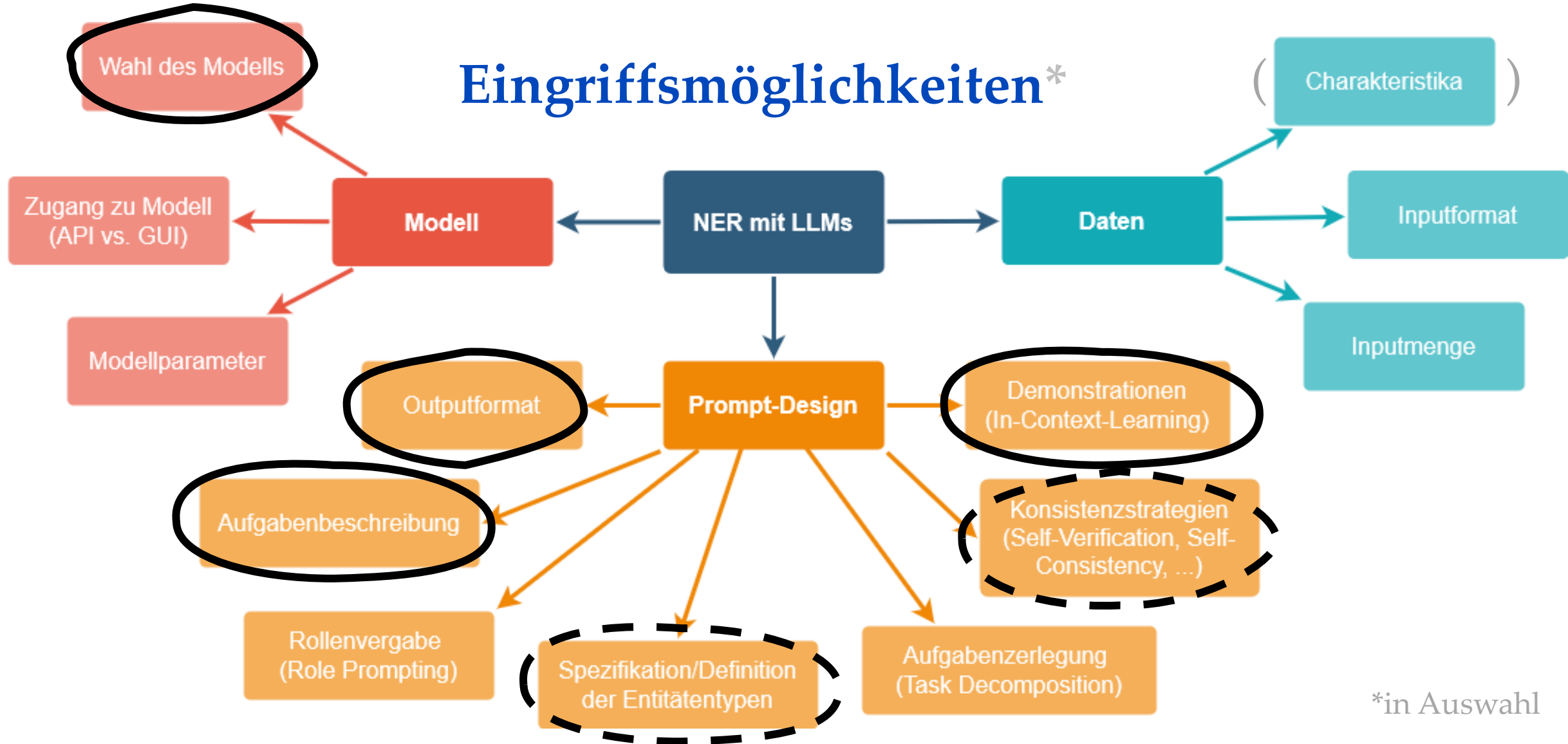


Eingriffsmöglichkeiten*



*in Auswahl

Eingriffsmöglichkeiten*



*in Auswahl

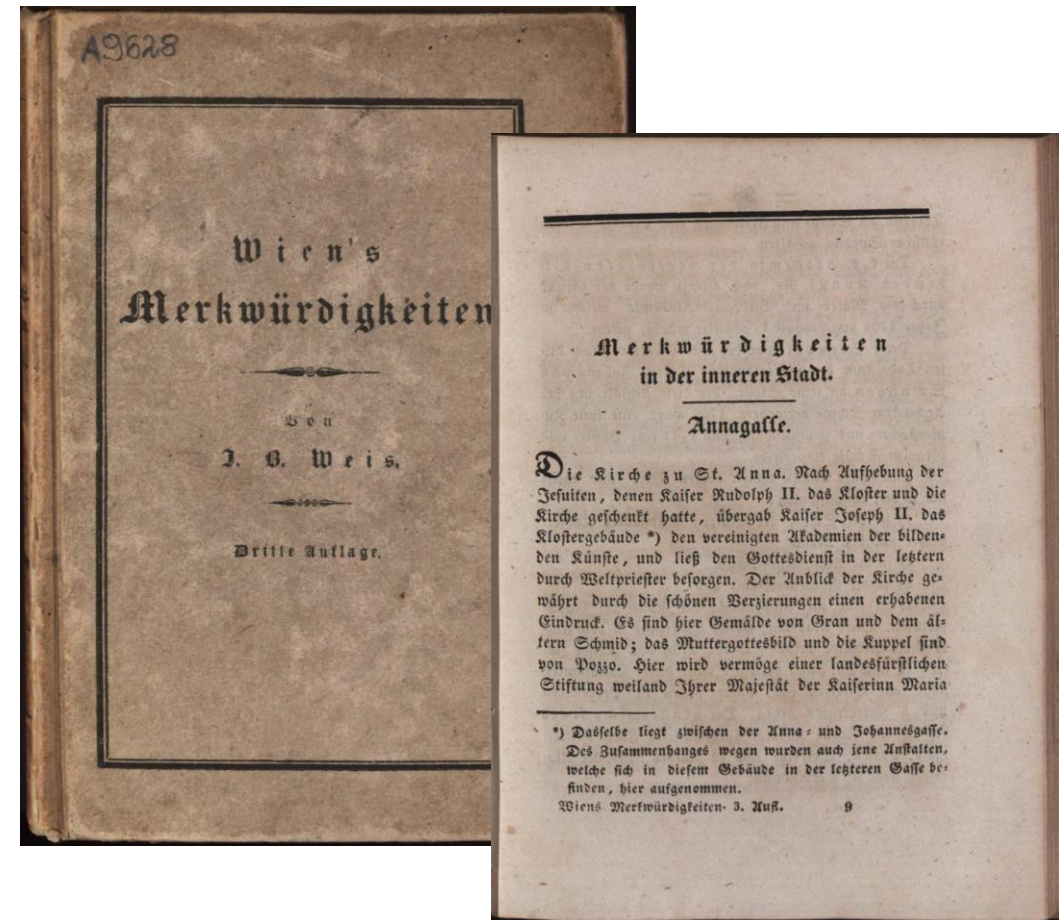
Wahl des Modells

- Eigene Use-Cases: GPT
- Neuere GPT-Versionen performen besser als ältere GPT-Versionen: $GPT-3 < GPT-3.5 < GPT-4$ (Sousa et al. 2023; Xie et al. 2023; Hu et al. 2023; Ashok & Lipton 2023, Covas 2023; Evans, Sadruddin & D'Souza 2024; García-Barragán et al. 2024; Hu et al. 2024; Munnangi et al. 2024)
- Eigene Use-Cases: GPT-3.5 und GPT-4o-mini > hohe NER-Scores auch mit kleineren/günstigeren Modellen erzielbar
- Vielzahl an Modellen verfügbar: GPT, Claude, Llama, Mixtral, Gemma, Grok, Gemini, ...



Experiment: Wahl des Modells

- Datengrundlage: Kapitel 4 des **Reisehandbuchs „Wien's Merkwürdigkeiten“** (Weiß 1836)
- Layout- und Texterkennung mithilfe von Transkribus (+ manuelle Überprüfung)
- Segmentierung in Sätze durch Tokenizer SoMaJo (Proisl & Uhrig 2016)
- 742 Sätze und 13.156 Tokens
- Ziel: Erkennung von Entitäten des Typs PERSON



Experiment: Wahl des Modells

You are an expert in Named Entity Recognition and Classification. role

Your task is to recognize named entities of the type PERSON in the given text. task

Return the entities in form of a list. output format

*Nach Aufhebung der Jesuiten, denen Kaiser Rudolph II. das Kloster und die Kirche geschenkt hatte, übergab Kaiser Joseph II. das Klostergebäude *) den vereinigten Akademien der bildenden Künste, und ließ den Gottesdienst in der letztern durch Weltpriester besorgen.*

['Kaiser Rudolph II. ', 'Kaiser Joseph II. ']

example 1

Der Anblick der Kirche gewährt durch die schönen Verzierungen einen erhabenen Eindruck.

[]

example 2

Es sind hier Gemälde von Gran und dem ältern Schmid; das Muttergottesbild und die Kuppel sind von Pozzo.

['Gran', 'ältern Schmid', 'Pozzo']

example 3



llama-3.1-70b-versatile
vs.
mixtral-8x7b-32768
vs.
gemma2-9b-it
vs.
gpt-4o-mini

```
[ ] def generate(input, model):

    chat_completion = client.chat.completions.create(
        messages=[
            {
                "role": "system",
                "content": """You are an expert in Named Entity Recognition and Classification.
                Your task is to recognize named entities of the type PERSON and in the given text.
                Return the entities in form of a list.

                Nach Aufhebung der Jesuiten, denen Kaiser Rudolph II. das Kloster und die Kirche geschenkt hatte, übergab
                Kaiser Joseph II. das Klostergebäude *) den vereinigten Akademien der bildenden Künste, und ließ den
                Gottesdienst in der letztern durch Weltpriester besorgen.
                ['Kaiser Rudolph II. ', 'Kaiser Joseph II. ']

                Der Anblick der Kirche gewährt durch die schönen Verzierungen einen erhabenen Eindruck.
                []

                Es sind hier Gemälde von Gran und dem ältern Schmid; das Muttergottesbild und die Kuppel sind von Pozzo.
                ['Gran', 'ältern Schmid', 'Pozzo']"""
            },
            {
                "role": "user",
                "content": input,
            }
        ],
        model=model
    )

    return(chat_completion.choices[0].message.content)
```

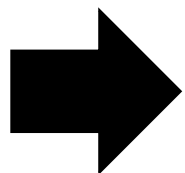
```
✓ [21] for x in df.index:
57 m print(x)
input = df.loc[x, "input"]
df.loc[x, "output_mixtral-8x7b-32768"] = generate(input, "mixtral-8x7b-32768")
```

	input
0	Merkwürdigkeiten in der inneren Stadt.
1	Annagasse.
2	Die Kirche zu St. Anna.
3	Nach Aufhebung der Jesuiten, denen Kaiser Rudo...
4	Der Anblick der Kirche gewährt durch die schön...

Ergebnisse: Wahl des Modells

Modell	llama-3.1-70b- versatile	mixtral-8x7b- 32768	gemma2-9b-it	gpt-4o-mini
Korrektes Outputformat (= valide Liste ohne Zusatz)	500 von 742 (67,4 %)	227 von 742 (30,6 %)	436 von 742 (58,8 %)	741 von 742 (99,9 %)
Recall*	0,88	0,64	0,81	0,82
Precision*	0,78	0,51	0,15	0,90
F1 Score*	0,83	0,57	0,25	0,86

* nur Sätze mit korrektem Outputformat einbezogen, partielle Matches erlaubt

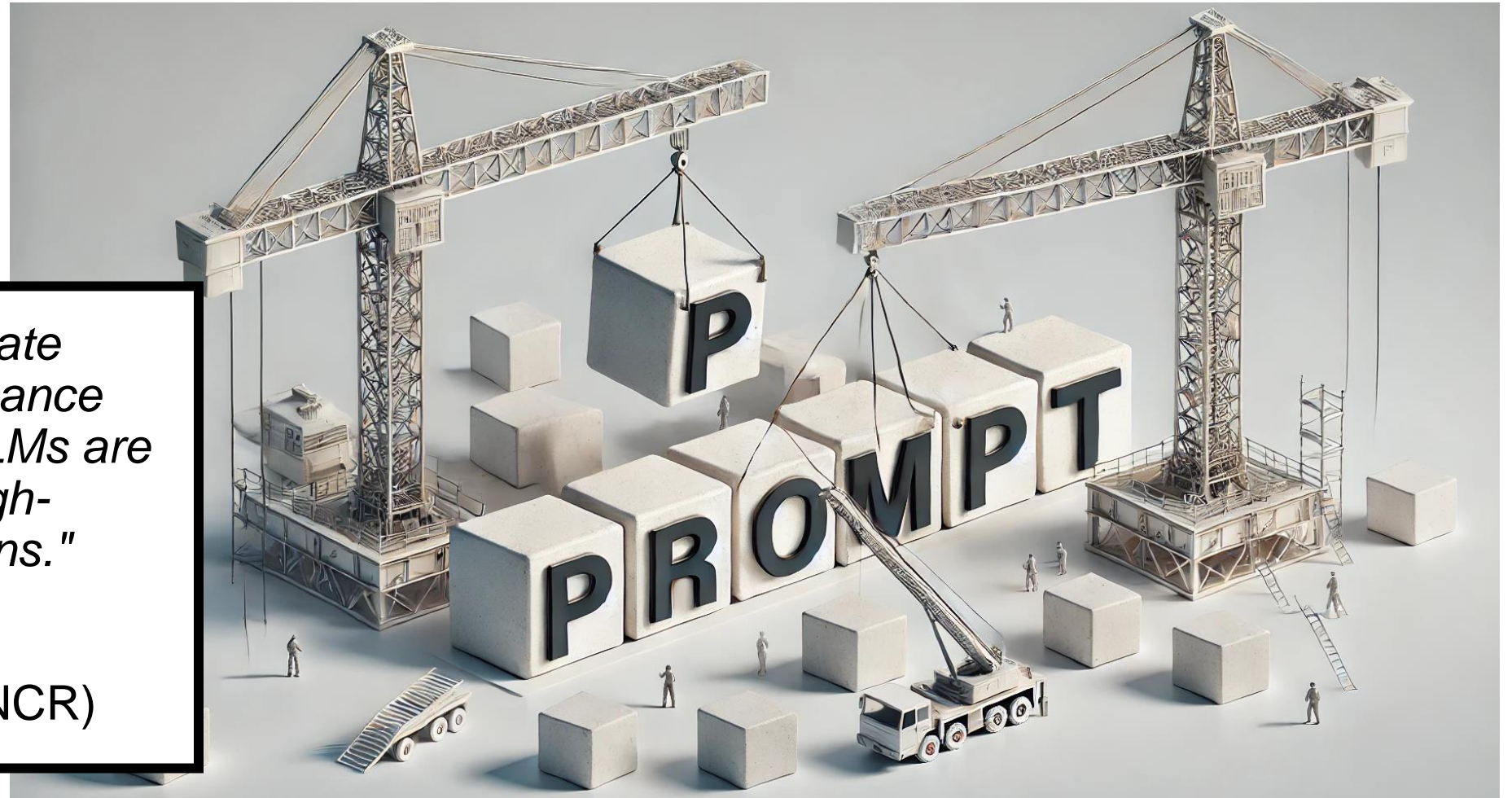


GPT-4o-mini schneidet (unter diesen vier Modellen) deutlich am besten ab

Von Modell zu Output

*"LLMs demonstrate superior performance [...] **when** the LLMs are provided with high-quality instructions."*

(Li & Zang 2024;
Hervorhebung: NCR)

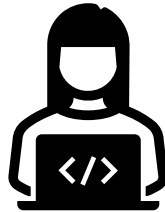


NER als Aufgabe für generative KI

"Extract without rephrasing all medical problem entities from the following note in a list format" (Hu et al. 2023)

"You want to perform a NER on a historical travelogue written in German language in the Year 1810." (Möhrke, Balck & Ananieva 2024)

"You take Passage as input and your task is to recognize and extract specific types of named entities in that given passage and classify into a set of entity types." (Rastinger 2024)



"What are the locations (LOC), persons (PER), organizations (ORG) and human productions (HumanProd) present in the following historical text?" (González-Gallardo et al. 2023)

„You have to perform a very specific task. Your primary task is to perform NER (Named Entity Recognition). So, your main task is to recognize and label 'Cancer Concept' entities using a specific format in which you must use @@ at the beginning and ## at the end of the 'Cancer Concept' entity." (García-Barragán et al. 2024)

"Please fulfill the following NER task by annotating the given scholarly paper title in the domain of astronomy." (Evans, Sadruddin & D'Souza 2024)

NER als Aufgabe für generative KI

- Wei et al. (2024): Vergleich von fünf verschiedenen NER-Prompts (alle anderen Faktoren konstant) → marginale Varianz von 0,003 %
- Genaue Formulierung nicht ausschlaggebend

ABER Aufgabe hier immer identisch:
Extraktion von Named Entities

No.	Template	F1(%)
1	Please recognize the entities of "" type in the given sentence: ""	45.08
2	Which entities of type "" are contained in the given sentence ""?	44.48
3	In the following sentence "", find the entities with type "".	45.12
4	Knowing the sentence "", identify the entities of type "" in it.	44.78
5	In the given sentence "", the entities of type "" are:	43.77
Average		44.65
Variance		0.003(%)

Wei et al. (2024): Table 4: Results on NER for different prompts

NER als Aufgabe für generative KI

- in bisherigen Studien und Experimenten drei „Paradigmen“ beobachtbar:

1) Extraktion

- Entitäten werden aus Inputtext extrahiert und (in strukturierter Form) ausgegeben
- Optional zusätzliche Informationen enthalten: Entitätentyp, Position, ...
- Outputformate: Liste, JSON, ...

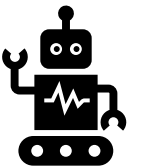


Die Burgkapelle, schon von Kaiser Friedrich III. im Jahre 1448 erbaut.

Input

Output

```
[{"Entität": „Burgkapelle“, „Label“: „LOC“}, {"Entität": „Kaiser Friedrich III.“, „Label“: „PER“}]
```



NER als Aufgabe für generative KI

- in bisherigen Studien und Experimenten drei „Paradigmen“ beobachtbar:

2) Annotation

- Gesamter Inputtext wird wiedergegeben, inklusive Markern vor und nach Entitäten
- Outputformate: „GPT-NER“ (= Auszeichnung über Sonderzeichen, Wang et al. 2023), HTML, ...

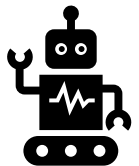


Die Burgkapelle, schon von Kaiser Friedrich III. im Jahre 1448 erbaut.

Input

Output

Die @@Burgkapelle##, schon von \$\$Kaiser Friedrich III.%% im Jahre 1448 erbaut.



NER als Aufgabe für generative KI

- in bisherigen Studien und Experimenten **drei „Paradigmen“** beobachtbar:

3) Transformation

- Inputtext wird in anderes Format überführt, in dem Entitäten basierend auf ihrer Position ausgezeichnet sind
- Outputformate: IOB/BIO, BMES, ...

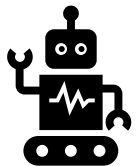


Die Burgkapelle, schon von Kaiser Friedrich III. im Jahre 1448 erbaut.

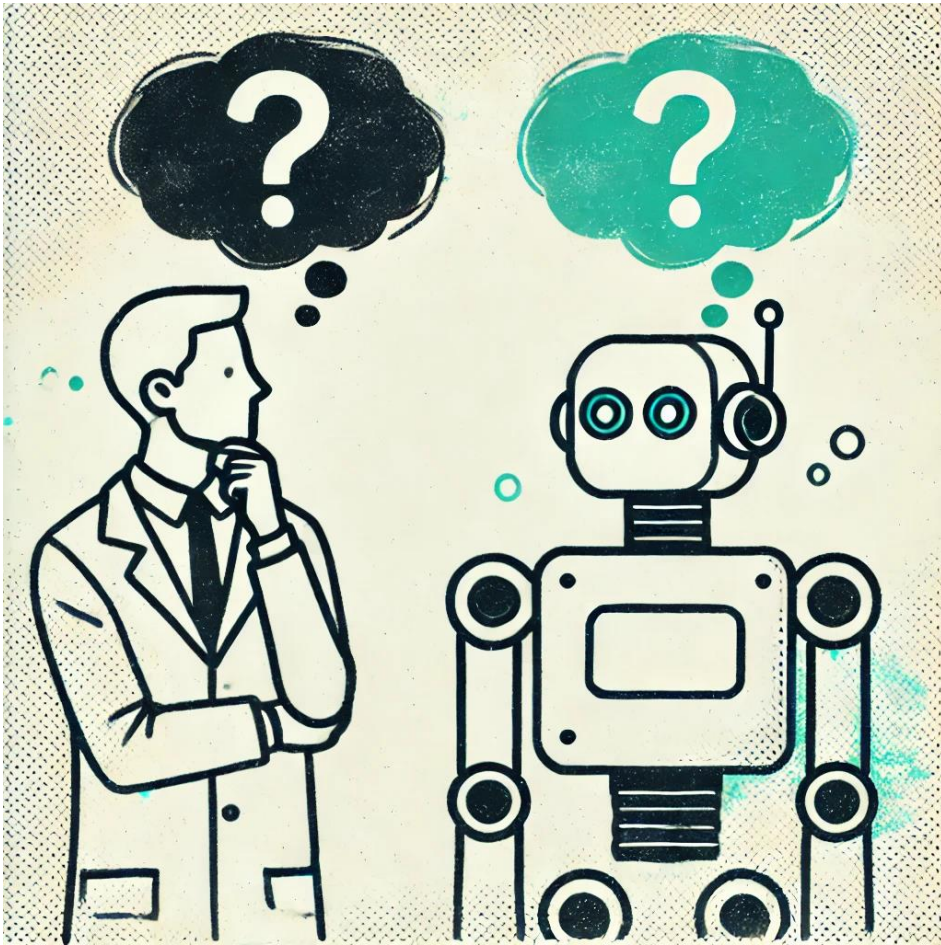
Input

Output

O B-LOC O O O B-PER I-PER I-PER I-PER O O O O O



NER: Extraktions-, Annotations- oder Transformationstask?



- Outputformate automatisiert ineinander überführbar > eigene Präferenzen sekundär
- Zentrale Frage: Welches Outputformat ist für LLMs am besten geeignet, d.h. erlaubt bestmögliches Ergebnis?

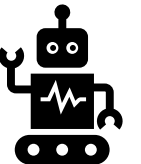
NER: Extraktion, Annotation oder Transformation?

- **Transformation** (IOB, BMES, ...):
inkonsistenter Output, niedrige F1-Scores > wenig geeignet (vgl. Wang et al. 2023, González-Gallardo et al. 2023)
- “it is usually even hard for the LLM to output a BMES string with the correct length, especially when the input sentence is long” (Wang et al. 2023)



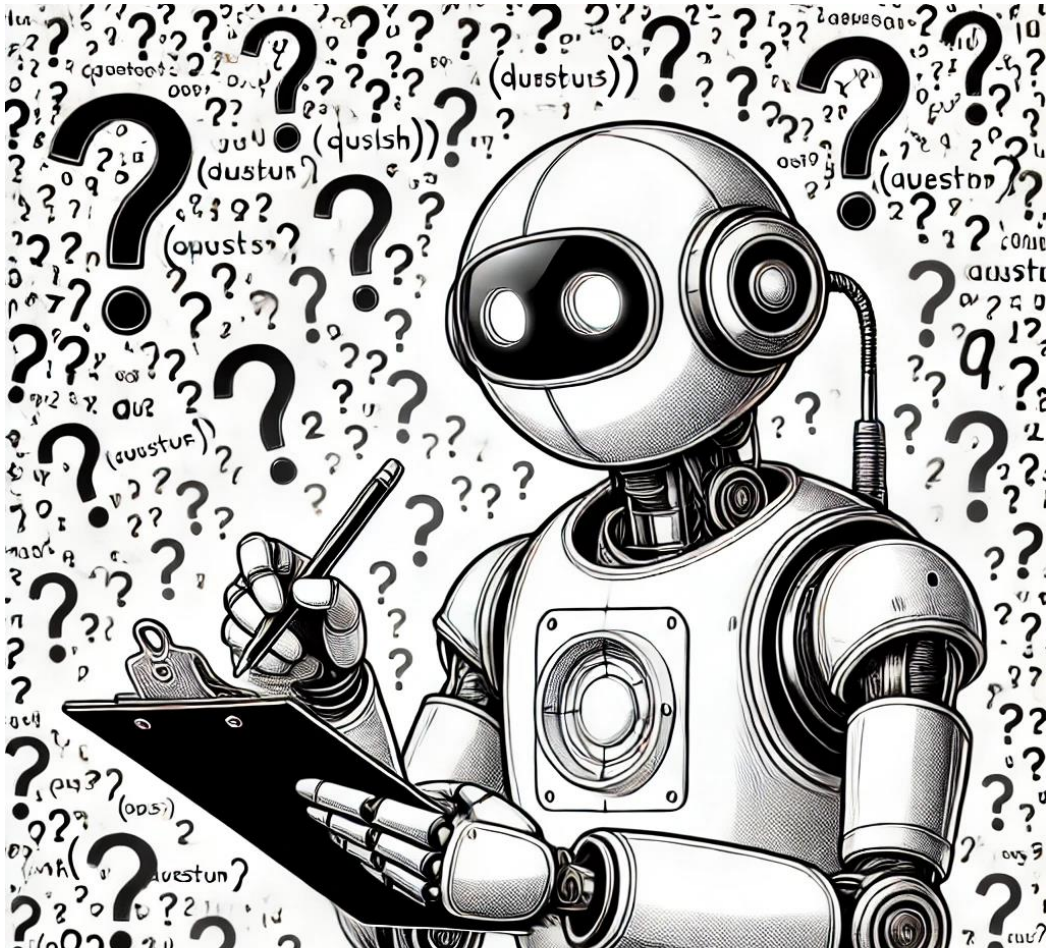
Die Burgstelle, schon zu Kaiser
Friedrich im Jahr 1180 erbaut.

O B-LOC O B-LOC I-PER I-
PER I-PER O O O O



→ weshalb?

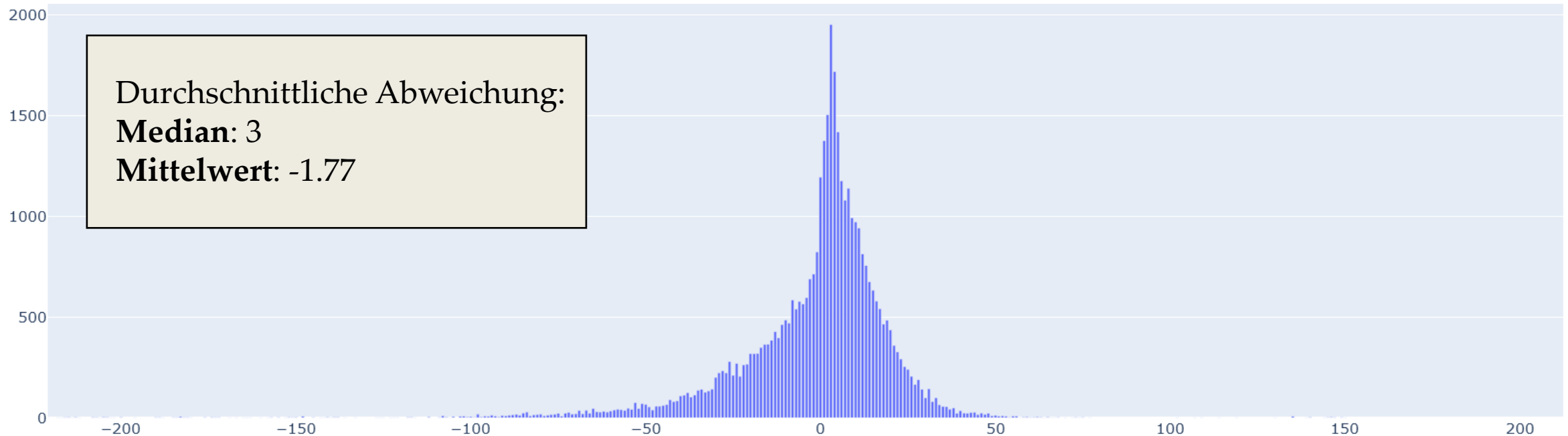
NER: Extraktion, Annotation oder ~~Transformation~~?



- Hintergrund: Zählen (von Zeichen, Tokens etc.) als Schwäche von LLMs
- (Optionale) Ausgabe von Positionsinformationen bei Extraktionstask ebenfalls nicht verlässlich
- vgl. Studie zu Ankunftslisten: Zeichenstart- und -endposition mit abgefragt

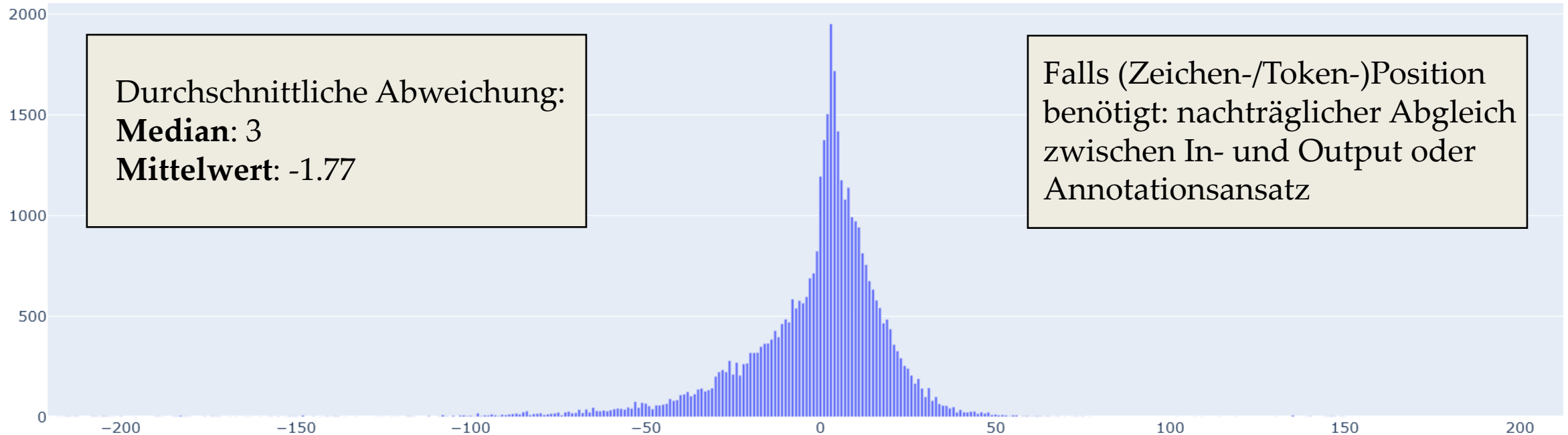
NER: Extraktion, Annotation oder ~~Transformation~~?

Differenz zwischen von GPT-3.5 geschätzter Startposition und tatsächlicher Startposition je Entität

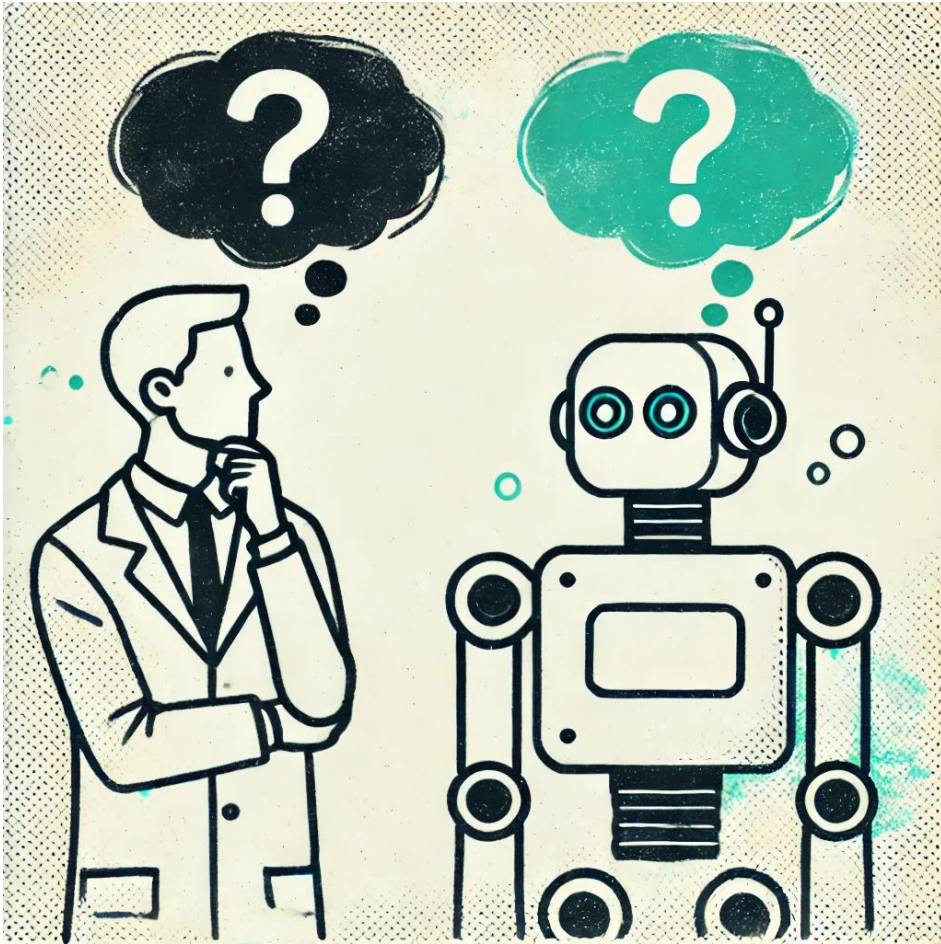


NER: Extraktion, Annotation oder ~~Transformation~~?

Differenz zwischen von GPT-3.5 geschätzter Startposition und tatsächlicher Startposition je Entität



NER: Extraktion, Annotation oder ~~Transformation~~?

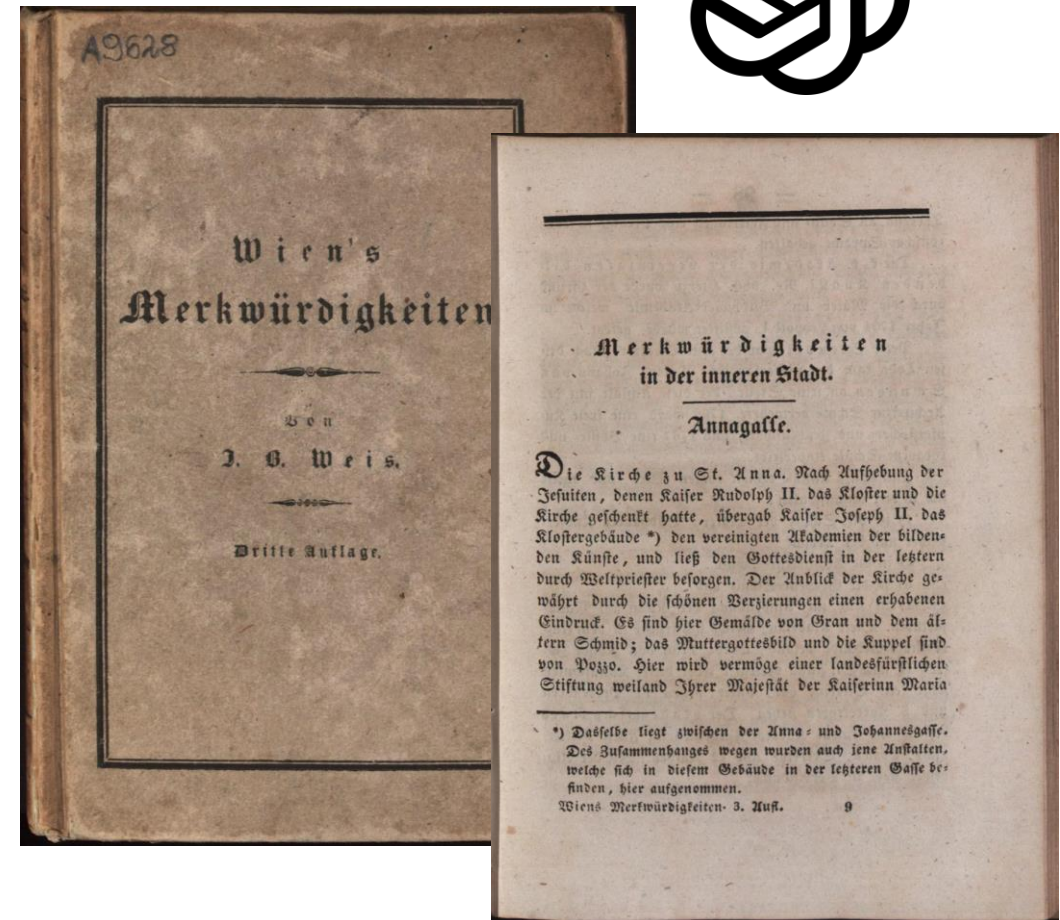


- Positionsbasierte Ansätze nicht zielführend
- Extraktion (ohne Position / Position nur als Schätzung) oder Annotation besser geeignet?
- beide Ansätze können sehr gute Ergebnisse liefern (vgl. z.B. zu Extraktion: Covas 2023, Wei et al. 2024, Mingchen et al. 2024, Rastinger 2024; zu Annotation: García-Barragán et al. 2024, Wang et al. 2023)
- Aber: bislang kein direkter empirischer Vergleich

Experiment: Extraktion vs. Annotation



- Gleiche Datengrundlage: Kapitel 4 des Reisehandbuchs „Wien's Merkwürdigkeiten“ (Weiß 1836)
- 6x NER über OpenAI API (GPT-4o-mini) durchgeführt – 3x Annotation / 3x Extraktion
- Self-Consistency-Strategie: Majority Voting (2/3) > höhere Reliabilität der Ergebnisse
- Alle anderen Parameter konstant gehalten



You are an expert in Named Entity Recognition and Classification. *role*
Your task is to recognize named entities of the type PERSON in the given text. *task*
Return the entities in form of a list. *output format*

*Nach Aufhebung der Jesuiten, denen Kaiser Rudolph II. das Kloster und die Kirche geschenkt hatte, übergab Kaiser Joseph II. das Klostergebäude *) den vereinigten Akademien der bildenden Künste, und ließ den Gottesdienst in der letztern durch Weltpriester besorgen.*

['Kaiser Rudolph II. ', 'Kaiser Joseph II. '] *example 1*

Der Anblick der Kirche gewährt durch die schönen Verzierungen einen erhabenen Eindruck.

[] *example 2*

Es sind hier Gemälde von Gran und dem ältern Schmid; das Muttergottesbild und die Kuppel sind von Pozzo.

['Gran', 'ältern Schmid', 'Pozzo'] *example 3*



Extraktion

You are an expert in Named Entity Recognition and Classification.

Your task is to recognize named entities of the type PERSON in the given text. task

Return the text with the entities marked with @@ at the beginning and ## at the end.

output format

*Nach Aufhebung der Jesuiten, denen Kaiser Rudolph II. das Kloster und die Kirche geschenkt hatte, übergab Kaiser Joseph II. das Klostergebäude *) den vereinigten Akademien der bildenden Künste, und ließ den Gottesdienst in der letztern durch Weltpriester besorgen.*

*Nach Aufhebung der Jesuiten, denen @@Kaiser Rudolph II.## das Kloster und die Kirche geschenkt hatte, übergab @@Kaiser Joseph II.## das Klostergebäude *) den vereinigten Akademien der bildenden Künste, und ließ den Gottesdienst in der letztern durch Weltpriester besorgen. example 1*

Der Anblick der Kirche gewährt durch die schönen Verzierungen einen erhabenen Eindruck.

Der Anblick der Kirche gewährt durch die schönen Verzierungen einen erhabenen Eindruck. example 2

Es sind hier Gemälde von Gran und dem ältern Schmid; das Muttergottesbild und die Kuppel sind von Pozzo.

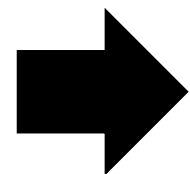
Es sind hier Gemälde von @@Gran## und dem @@ältern Schmid##; das Muttergottesbild und die Kuppel sind von @@Pozzo##. example 3



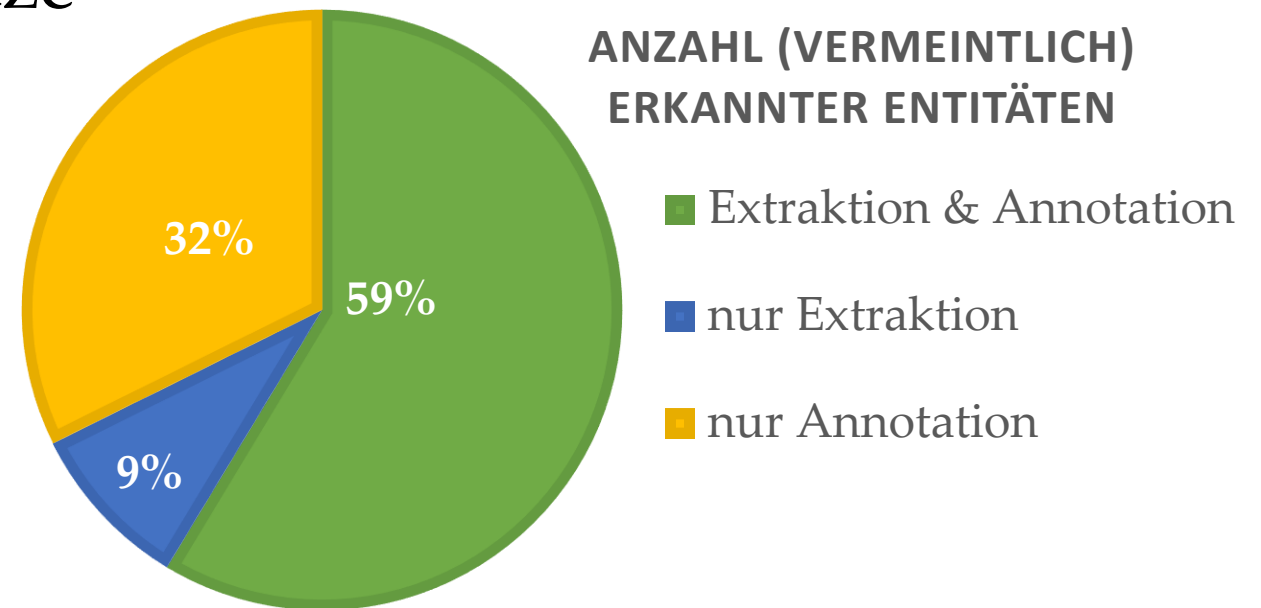
Annotation

Resultat: Annotation vs. Extraktion

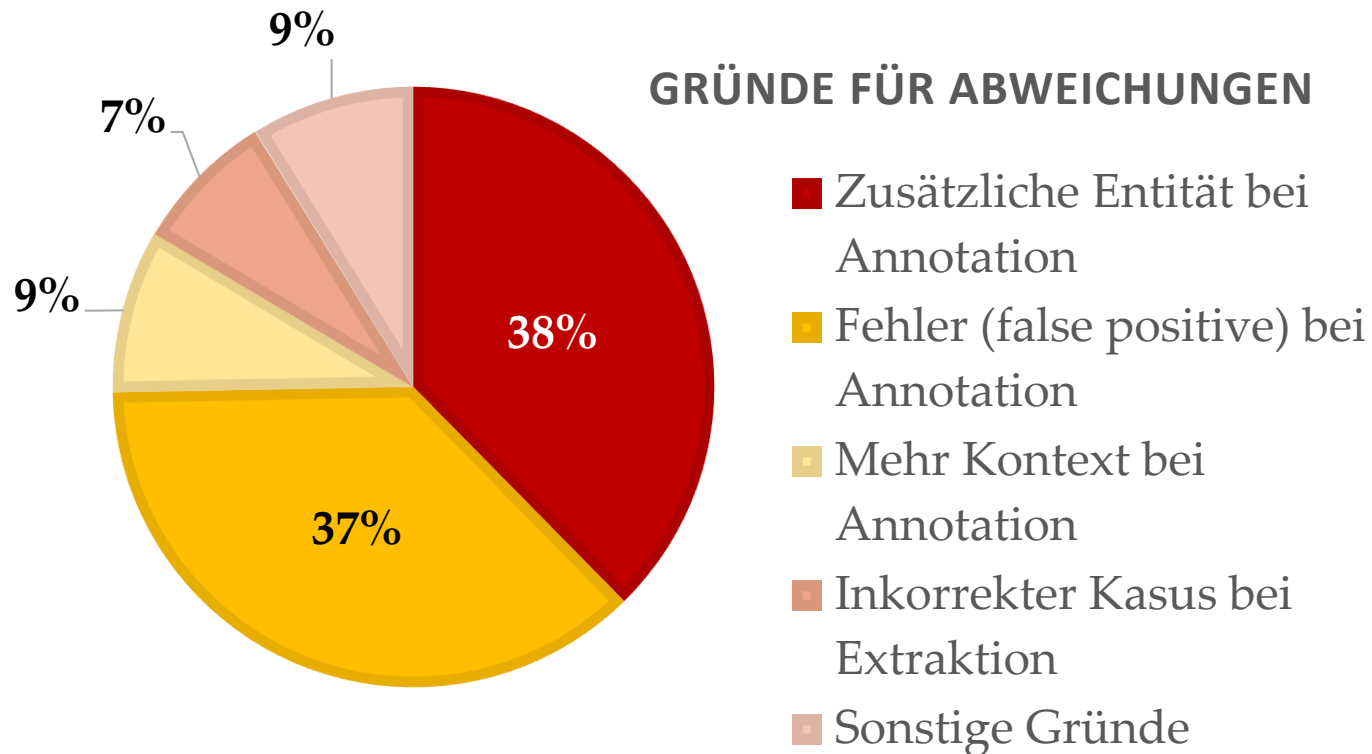
- 606 von 742 Sätze (81,7 %):
identische Entitäten > für 136 Sätze
unterschiedliche Ergebnisse
- Annotationsansatz ‚erkennt‘
wesentlich mehr Entitäten
- Quantität ≠ Qualität



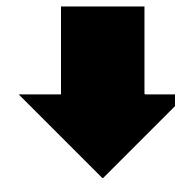
Qualitative Analyse
der Abweichungen



Resultat: Annotation vs. Extraktion



- **Entitätenmenge & -länge:** LLM [GPT-4o-mini] annotiert großzügiger als es extrahiert
- Mehr korrekte Entitäten (höherer Recall), aber auch mehr Fehler (niedrigere Precision)



Wahl abhängig von spezifischem Anwendungsszenario

Entscheidungshilfe: Annotation vs. Extraktion

1) Die ausgegebenen Entitäten

- a) sollen in Schreibung und grammatischer Form möglichst nah am Originaltext sein.
- b) können in Schreibung und grammatischer Form teilweise normalisiert sein.

2) Entitäten sollen möglichst ...

- a) großzügig erkannt werden, d.h. möglichst inklusive allem zugehörigen Kontext zurückgegeben werden.
- b) zurückhaltend erkannt werden, d.h. möglichst nur das Wesentliche zurückgegeben werden.

3) Besonders wichtig für den Erfolg der NER ist ...

- a) ein hoher Recall, d.h. dass möglichst viele Entitäten gefunden werden.
- b) eine hohe Precision, d.h. dass möglichst viele der erkannten Entitäten korrekt sind.

4) Die API-Kosten ...

- a) dürfen auch höher als unbedingt nötig ausfallen.
- b) sollten so gering wie möglich gehalten werden.

Entscheidungshilfe: Annotation vs. Extraktion

1) Die ausgegebenen Entitäten

- a) sollen in Schreibung und grammatischer Form möglichst nah am Originaltext sein.
- b) können in Schreibung und grammatischer Form teilweise normalisiert sein.

2) Entitäten sollen möglichst ...

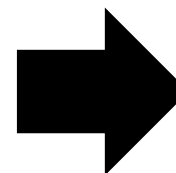
- a) großzügig erkannt werden, d.h. möglichst inklusive allem zugehörigen Kontext zurückgegeben werden.
- b) zurückhaltend erkannt werden, d.h. möglichst nur das Wesentliche zurückgegeben werden.

3) Besonders wichtig für den Erfolg der NER ist ...

- a) ein hoher Recall, d.h. dass möglichst viele Entitäten gefunden werden.
- b) eine hohe Precision, d.h. dass möglichst viele der erkannten Entitäten korrekt sind.

4) Die API-Kosten ...

- a) dürfen auch höher als unbedingt nötig ausfallen.
- b) sollten so gering wie möglich gehalten werden.



Überwiegend a) → Annotation
Überwiegend b) → Extraktion

Demonstrationen

You are an expert in Named Entity Recognition and Classification.

Your task is to recognize named entities of the type PERSON and in the given text. **task**

Return the text with the entities marked with @@ at the beginning and ## at the end.

output format

Nach Aufhebung der Jesuiten, denen Kaiser Rudolph II. das Kloster und die Kirche geschenkt hatte, übergab Kaiser Joseph II. das Klostergebäude *) den vereinigten Akademien der bildenden Künste, und ließ den Gottesdienst in der letztern durch Weltpriester besorgen.

Nach Aufhebung der Jesuiten, denen @@Kaiser Rudolph II.## das Kloster und die Kirche geschenkt hatte, übergab @@Kaiser Joseph II.## das Klostergebäude *) den vereinigten Akademien der bildenden Künste, und ließ den Gottesdienst in der letztern durch Weltpriester besorgen.

example 1

Der Anblick der Kirche gewährt durch die schönen Verzierungen einen erhabenen Eindruck.

Der Anblick der Kirche gewährt durch die schönen Verzierungen einen erhabenen Eindruck.

example 2

Es sind hier Gemälde von Gran und dem ältern Schmid; das Muttergottesbild und die Kuppel sind von Pozzo.

Es sind hier Gemälde von @@Gran## und dem @@ältern Schmid##; das Muttergottesbild und die Kuppel sind von @@Pozzo##.

example 3

You are an expert in Named Entity Recognition and Classification.
Your task is to recognize named entities of the type PERSON and in the given text.
Return the entities in form of a list.

Nach Aufhebung der Jesuiten, denen Kaiser Rudolph II. das Kloster und die Kirche geschenkt hatte, übergab Kaiser Joseph II. das Klostergebäude *) den vereinigten Akademien der bildenden Künste, und ließ den Gottesdienst in der letztern durch Weltpriester besorgen.

['Kaiser Rudolph II.', 'Kaiser Joseph II. ']

Der Anblick der Kirche gewährt durch die schönen Verzierungen einen erhabenen Eindruck.

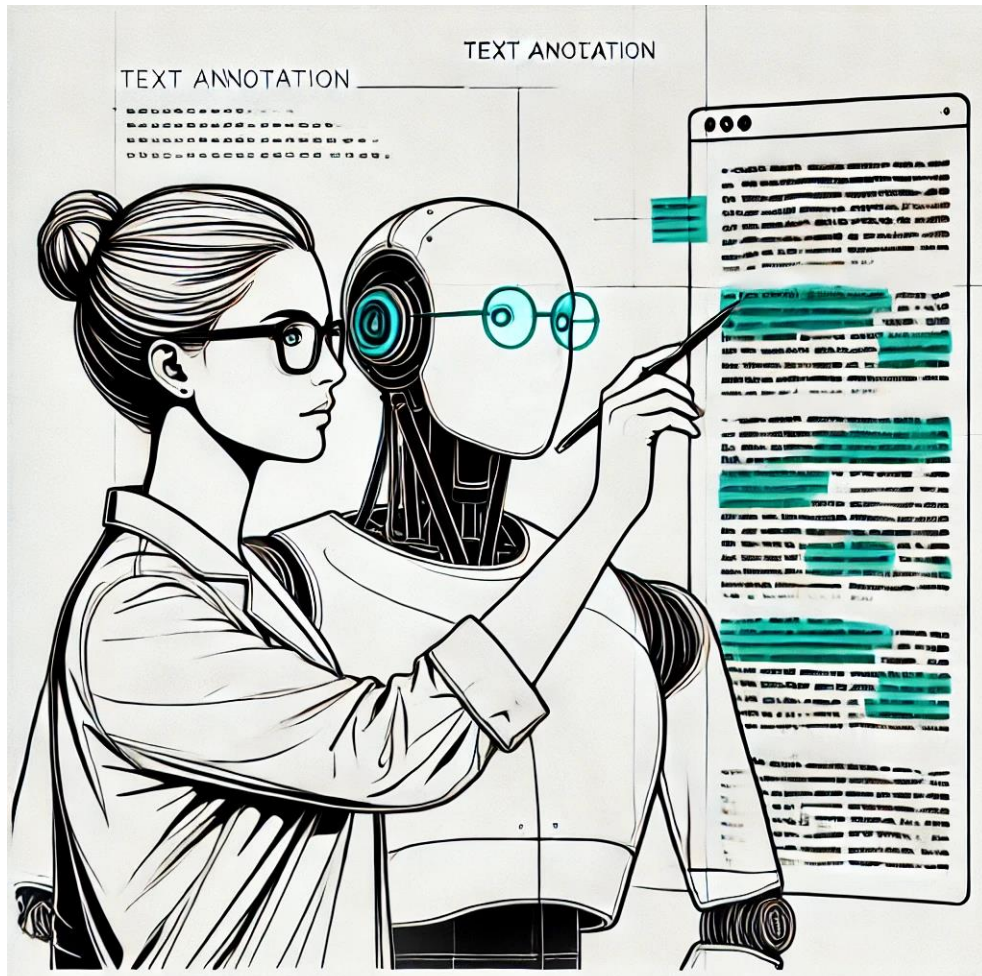
[]

Es sind hier Gemälde von Gran und dem ältern Schmid; das Muttergottesbild und die Kuppel sind von Pozzo.

['Gran', 'ältern Schmid', 'Pozzo']

example 3

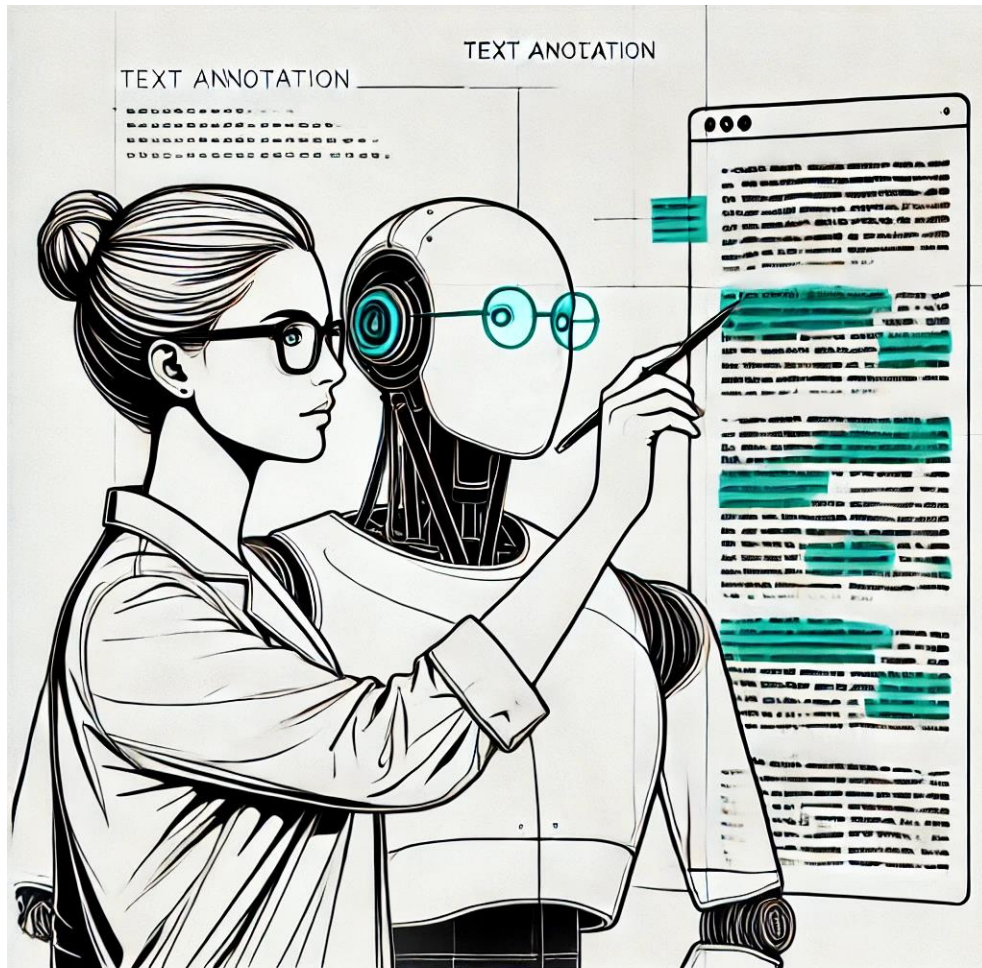
Demonstrationen („In-Context-Learning“)



- Beispielhafte Input-Output-Paare (“Demonstrationen”)
- Differenzierung dreier Settings:
 - 1) **Zero-Shot**: keine Beispiele
 - 2) **One-Shot**: ein Beispiel
 - 3) **Few-Shot**: mindestens zwei Beispiele

→ Sollte man LLMs NER demonstrieren?

Demonstrationen („In-Context-Learning“)



→ Sollte man LLMs NER demonstrieren? Ja!

- Persönliche Erfahrung: One- oder Few-Shot-Settings liefern wesentlich bessere Ergebnisse
- Gleiches auch in Vielzahl an Studien belegt (z.B. Ashok & Lipton 2023; Chantran et al. 2023; Covas 2023; García-Barragán et al. 2024; Hu et al. 2024; Munnangi et al. 2024; Wang et al. 2023; Wei et al. 2024; Xie et al. 2023)
- Nachteile: höhere Kosten (minimal, da Inputtokens), zusätzlicher Aufwand (minimal, da geringe Anzahl oft bereits ausreichend)
- Vorteile: schnelle Anpassung an eigenes Datenmaterial, konsistenterer Output, Alternative zu langen Annotationsrichtlinien

Demonstrationen („In-Context-Learning“)

Ankunfft Hoher und Nidriger Stands=Persohnen.

Rothen Thurn / den 15. Aug. Herr Obrist=Leuthen. Gundager vom Mercischen Regiment / kombt auß dem Reich / logirt in der gulden Sonn. Herr Hauptmann Kayser / vom Marsiglischen Regiment / kombt von Breyfach / logirt im gulden Hirschen. Stuben=Thor / Jhro Hochwürden Herr Emericus Nyllas, Thomherr in Agramb / neuer Rector deß Croatischen Collegii allhier / kombt auß Croathen / und logirt in besagtem Collegio.

Cärnter Thor / den 17. Augusti, Ein Currier / kommet zur Post von Düsseldorff / logirt im Oelerischen Hauß / in der Cärner=Strassen.

PERSON

STADTTOR

DATUM

HERKUNFTSORT

UNTERKUNFTSORT

ZIELORT



Demonstrationen („In-Context-Learning“)

Ankunfft Hoher und Nidriger Stands=Persohnen.

Rothen Thurn / den 15. Aug. Herr Obrist=Leuthen. Gundager vom
Mercischen Regiment / kombt auß dem Reich / logirt in der gulden
Sonn. Herr Hauptmann Kayser / vom Marsiglichen Regiment / kombt von
Breyfach / logirt im gulden Hirschen. Stuben=Thor / Jhro Hochwürden
Herr Emericus Nyllas, Thomherr in Agramb / neuer Rector deß
Croatischen Collegii allhier / kombt auß Croathen / und logirt in
besagtem Collegio.

Cärnter Thor / den 17. Augusti, Ein Currier / kommet zur Post von
Düsseldorff / logirt im Oelerischen Hauß / in der Cärner=Strassen.

PERSON STADTTOR DATUM HERKUNFTSORT UNTERKUNFTSORT ZIELORT

Zero-Shot +
Spezifikation der
Entitätentypen

Demonstrationen („In-Context-Learning“)

Ankunft Hoher und Nidriger Stands=Persohnen.

Rothen Thurn / den 15. Aug. Herr Obrist=Leuthen. Gundager vom
Mercischen Regiment / kombt auß dem Reich / logirt in der gulden
Sonn. Herr Hauptmann Kayser / vom Marsiglichen Regiment / kombt von
Breyfach / logirt im gulden Hirschen. Stuben=Thor / Jhro Hochwürden
Herr Emericus Nyllas, Thomherr in Agramb / neuer Rector deß
Croatischen Collegii allhier / kombt auß Croathen / und logirt in
besagtem Collegio.

Cärnter Thor / den 17. Augusti, Ein Currier / kommet zur Post von
Düsseldorff / logirt im Oelerischen Hauß / in der Cärner=Strassen.

PERSON STADTTOR DATUM HERKUNFTSORT UNTERKUNFTSORT ZIELORT

Zero- > One-
bzw. Few-Shot
(1 Liste, 9 Items)

Use Case: Ankunftslisten des *Wienerischen Diariums*

39.500 Entitäten aus 1.447
Ankunftslisten (1703-
1725) extrahiert

Evaluation:

Precision: 0.98

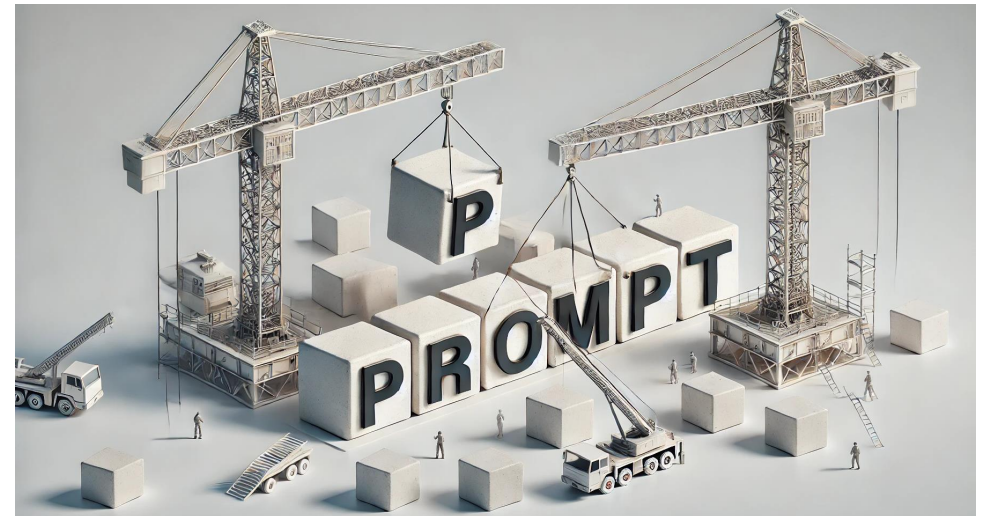
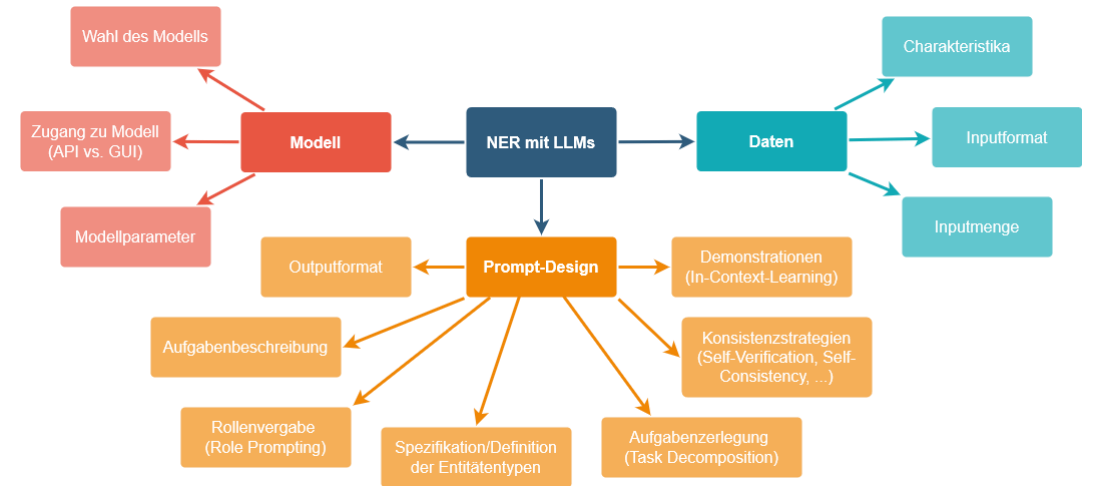
Recall: 0.97

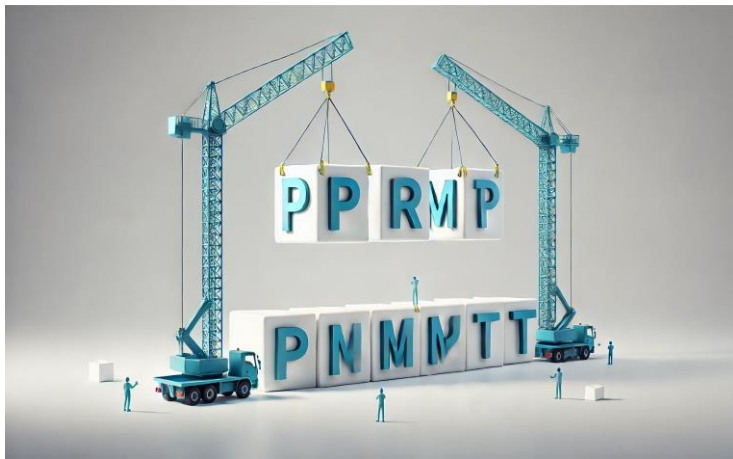
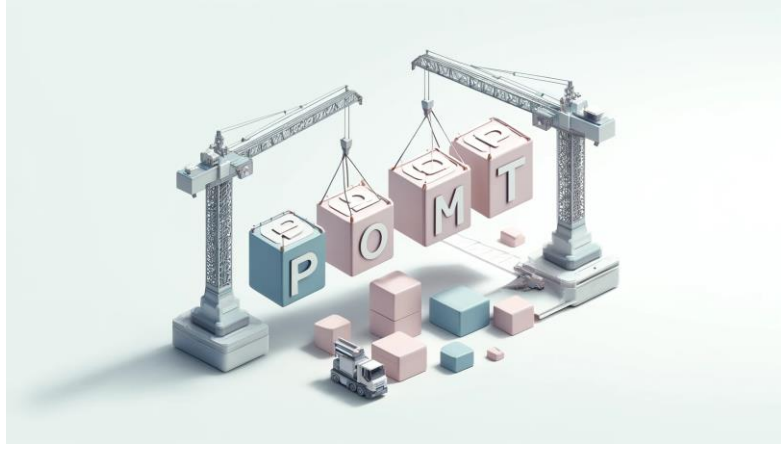
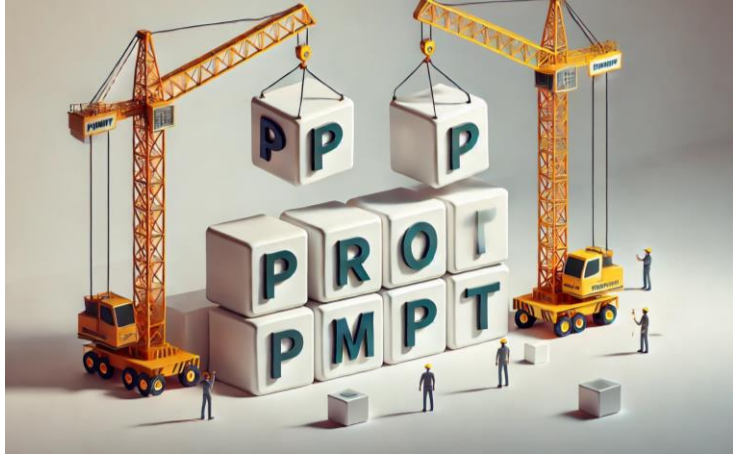
F1-Score: 0.97

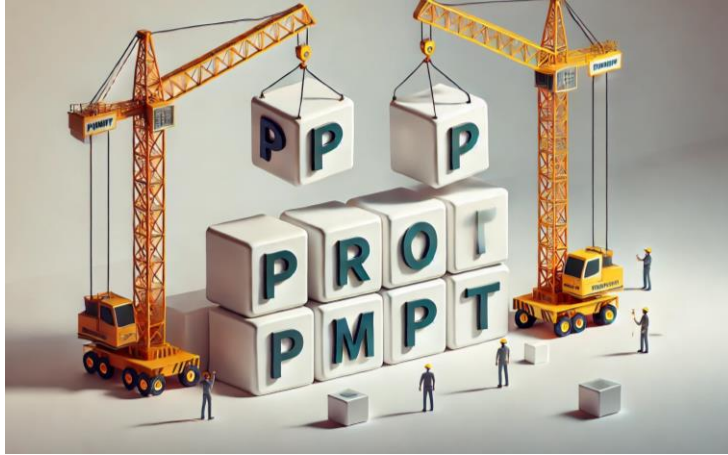


(Zwischen-)Fazit

- Vielzahl an Möglichkeiten NER mit LLMs an eigene Daten anzupassen
- Weitere systematische und groß angelegte Experimente (mit historischen Texten) notwendig
- Scheitern als Fortschritt, Erkenntnis als iterativer Prozess (z.B. Transformation)
- vgl. Bildgenerierung mit Dall-E: viele Fehlschritte bis zum richtigen „PROMPT“







Vielen Dank für die Aufmerksamkeit!



Referenzen

- Bilder generiert mit DALL-E
- Ashok, Dhananjay & Zachary C. Lipton (2023): PromptNER: Prompting for Named Entity Recognition. <https://doi.org/10.48550/arXiv.2305.15444>
- Chantran, Mohan Raj, Lay-Ki Soon, Huey Fang Ong, & Bhawani Selvaretnam (2023): How well ChatGPT understand Malaysian English? An Evaluation on Named Entity Recognition and Relation Extraction. <https://doi.org/10.48550/arXiv.2311.11583>
- Covas, Eurico (2023): Named entity recognition using GPT for identifying comparable companies. <https://doi.org/10.48550/arXiv.2307.07420>
- Evans, Julia, Sameer Sadruddin, & Jennifer D'Souza (2024): Astro-NER – Astronomy Named Entity Recognition: Is GPT a Good Domain Expert Annotator? <https://doi.org/10.48550/arXiv.2405.02602>
- García-Barragán, Álvaro et al. (2024): GPT for medical entity recognition in Spanish. <https://doi.org/10.1007/s11042-024-19209-5>
- González-Gallardo, Carlos-Emiliano et al. (2023): Yes but.. Can ChatGPT Identify Entities in Historical Documents? <https://doi.org/10.48550/arXiv.2303.17322>
- Hu, Yan et al. (2023): Zero-shot clinical entity recognition using ChatGPT. <https://doi.org/10.48550/arXiv.2303.16416>
- Hu, Yan et al. (2024): Improving large language models for clinical named entity recognition via prompt engineering. <https://doi.org/10.1093/jamia/ocad259>

Referenzen

- Mingchen, Li, & Rui Zhang (2023): How far is Language Model from 100% Few-shot Named Entity Recognition in Medical Domain. <https://doi.org/10.48550/arXiv.2307.00186>
- Mingchen, Li, Huixue Zhou, Han Yang, Rui Zhang (2024): RT: a Retrieving and Chain-of-Thought framework for few-shot medical named entity recognition. <https://doi.org/10.1093/jamia/ocae095>
- Möhrke, Jacob, Sandra Balck, & Anna Ananieva (2024): Zum Einsatz von GPT 4 für NER: Ein Experiment anhand eines historischen Reisetextes. <https://zenodo.org/doi/10.5281/zenodo.10843214>
- Munnangi, Monica et al.(2024): On-the-fly Definition Augmentation of LLMs for Biomedical NER. <https://doi.org/10.48550/arXiv.2404.00152>
- Rastinger, Nina C. (2024): Re-Reading Lists in Historical Newspapers: Digital Insights into an Overlooked Text Type. <https://doi.org/10.3384/ecp210016>
- Sousa, Hugo, Nuno Guimarães, Alípio Jorge, & Ricardo Campos (2023): GPT Struct Me: Probing GPT Models on Narrative Entity Extraction. <https://doi.org/10.48550/arXiv.2311.14583>
- Tsitseklis, Konstantinos, Georgia Stavropoulou, & Symeon Papavassiliou (2024): Custom Named Entity Recognition VS ChatGPT Prompting: A Paleontology Experiment. <https://doi.org/10.1109/PACET60398.2024.10497008>

Referenzen

- Wang, Shuhe et al. (2023): GPT-NER: Named Entity Recognition via Large Language Models. <https://doi.org/10.48550/arXiv.2304.10428>
- Wei, Xiang, Xingyu Cui, Ning Cheng, Xiaobin Wang, Xin Zhang, Shen Huang, Pengjun Xie, Jinan Xu, Yufeng Chen, Meishan Zhang, Yong Jiang, & Wenjuan Han (2024): ChatIE: Zero-Shot Information Extraction via Chatting with ChatGPT. <https://doi.org/10.48550/arXiv.2302.10205>
- Xie, Tingyu, Qi Li, Jian Zhang, Yan Zhang, Zuozhu Liu, & Hongwei Wang (2023): Empirical Study of Zero-Shot NER with ChatGPT. <https://doi.org/10.48550/arXiv.2310.10035>
- Xie, Tingyu, Qi Li, Yan Zhang, Zuozhu Liu, Hongwei Wang (2024): Self-Improving for Zero-Shot Named Entity Recognition with Large Language Models. <https://doi.org/10.48550/arXiv.2311.08921>