



# Zum Einsatz von GPT-4 für NER: Ein Experiment anhand eines historischen Reisetextes

Jacob Möhrke, Sandra Balck & Anna  
Ananieva

Leibniz-Institut für Ost- und  
Südosteuropaforschung

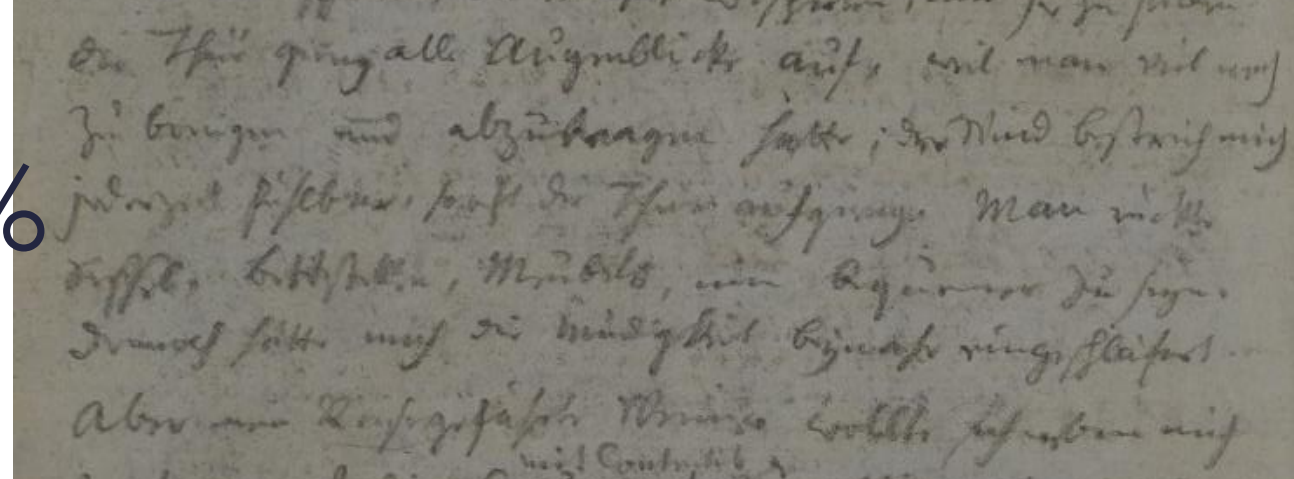
DFG-Projekt "Digitale Editionen historischer  
Reiseberichte" (DEHisRe)

# @@Abstract##

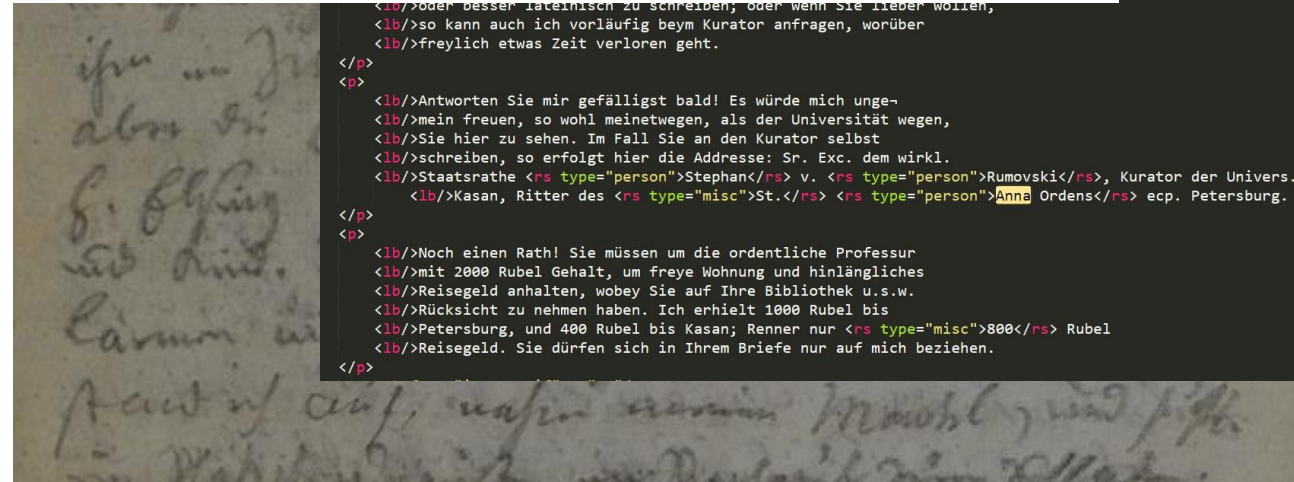
- > Historische Texte sind für NER besonders herausfordernd (Ehrmann et al)
  - > Large Language Models (LLMs) können für historische Texte effektiver sein als herkömmliche Modelle (De Toni et al)
  - > LLMs sind auf Textgenerierung ausgelegt
  - > NER ist eine Sequenzierungsaufgabe
- ⇒ wir wollen eine Möglichkeit präsentieren diese Lücken zu überbrücken

# ~~Ausgangslage%%

- > Quelltext: Reisebericht von F. X. Bronner 1810 (1817 Rückreise)
- > Experiment an einem Ausschnitt von ca. 6000 Wörtern
- > Händisch ausgezeichnet 295 Entitäten (Person, Ort, Organisation, Datum, Werk, Währung und Sonstiges)  
⇒ für Evaluation
- > Ungenügende Resultate aus herkömmlichen NER-Ansätzen (spacy, flair, etc.)



```
<lb/>vielleicht eine Art Lehrbuch liegen haben, so würde die Sache  
<lb/>keine Schwierigkeit finden. Im letztern Falle wäre es vielleicht  
<lb/>am zweckmäßigsten, unmittelbar an den <rs type="person" facts="neu_fxb_69">Kurator</rs>, deutsch  
<lb/>oder besser lateinisch zu schreiben; oder wenn Sie lieber wollen,  
<lb/>so kann auch ich vorläufig beym <rs type="person" facts="neu_fxb_69">Kurator</rs> anfragen, worüber  
<lb/>freylich etwas Zeit verloren geht.  
</p>  
<p>  
<lb/>Antworten Sie mir gefälligst bald! Es würde mich unge-  
<lb/>mein freuen, so wohl meinetwegen, als der <rs type="org">Universität</rs> wegen,  
<lb/>Sie hier zu sehen. Im Fall Sie an den <rs type="person" facts="neu_fxb_69">Kurator</rs> selbst  
<lb/>schreiben, so erfolgt hier die Adresse: Sr. Exc. dem wirkl.  
<lb/>Staatsrathe Stephan v. <rs type="person" facts="neu_fxb_69">Rumovski</rs>, Kurator der <rs type="org">Univers.  
<lb/>Kasan</rs>, Ritter des <rs type="org">St. Anna Ordens</rs> ecp. Petersburg.  
</p>  
<p>  
<lb/>Noch einen Rath! Sie müssen um die ordentliche Professur  
<lb/>mit <rs type="currency">2000 Rubel</rs> Gehalt, um freye Wohnung und hinlängliches  
<lb/>Reisegeld anhalten, wobey Sie auf Ihre Bibliothek u.s.w.  
<lb/>Rücksicht zu nehmen haben. Ich erhielt <rs type="currency">1000 Rubel</rs> bis  
<lb/><rs type="place">Petersburg</rs>, und <rs type="currency">400 Rubel</rs> bis <rs type="place">Kasan</rs>; <rs  
<lb/>Reisegeld. Sie dürfen sich in Ihrem Briefe nur auf mich beziehen.  
</p>
```



# %%GPT-NER (Wang et al)&&

- > Transformation der "Sequence Labeling Task" NER in "Generation Task" (GPT & andere LLMs)
- > Verwendung von "Speziellen Tokens" (z.B. @@ ##) zum Markieren der Entitäten
- > Few-Shot Ansatz
- > "Halluzination" von GPT als Problem

# %%GPT-NER (Wang et al)&&

I am an excellent linguist. The task is to label location entities in the given sentence. Below are some examples

Task Description

Input: Only France and Britain backed Fischler 's proposal . **Example 1**

Output: Only @@France## and @@Britain## backed Fischler 's proposal .

Input: Germany imported 47,600 sheep from Britain last year , nearly half of total imports . **Example 2**

Output: @@Germany## imported 47,600 sheep from @@Britain## last year , nearly half of total imports .

Few-shot Demonstrations

Input: It brought in 4275 tonnes of British mutton . some 10 percent of overall imports . **Example 3**

Output: It brought in 4275 tonnes of British mutton . some 10 percent of overall imports .

Input: China says Taiwan spoils atmosphere for talks .

Output: @@China## says @@Taiwan## spoils atmosphere for talks .

Input Sentence

Beispiel eines Instruktions-Prompts von Wang et al (S. 4)

# §§ Experiment~~

> Custom-GPT:

<https://chat.openai.com/g/g-suHf2Xhom-dh-assistant>

> Auffinden und Klassifizieren der Entitäten:

<https://chat.openai.com/share/79709912-f615-4e12-8c51-f50aeb5c3325>

> Versuch der Normalisierung:

<https://chat.openai.com/share/c3a2cf75-15eb-4cfc-a5b1-b05564e4d141>



## DH Assistant

By Jacob Moehrke

Expert in Digital Humanities, assisting with code formatting, data curation, TEI-XML, and NLP tasks.

How do I format this code for better readability?

I need assistance with TEI-XML formatting.

Can you help me curate this dataset?

How do I use NER in my research?

 Message DH Assistant...



ChatGPT can make mistakes. Consider checking important information.

# §§ Experiment ~ ~ NER



## Anonymous

You want to perform a NER on a historical travelogue written in German language in the Year 1810. The entities you want to detect are "Person", "Location", "Organisation", "Date", "Currency", and "Work" (e.g. mainly writings). You want to do it using only GPT. In order to do so it is your task to identify and label all entities in the given texts, with the use of a specific format. Use the format '##Entity++' for Person, '@@Entity##' for Location, '~~Entity%%' for Date, '++Entity~~' for Organisation, '%%Entity&&' for Currency, and '\$\$Entity~~' for Work. Take a deep breath and work on this problem step-by-step.

# §§ Experiment ~ ~ NER



## **DH Assistant**

To approach this task, we'll need to proceed step-by-step, as it involves several complex elements: understanding the text, identifying specific entities, and applying a unique labeling format. Given that I am unable to directly read or process external texts, you will need to provide the text from the historical travelogue.

Once you provide the text, I will manually analyze it to identify entities in the categories of "Person", "Location", "Organisation", "Date", "Currency", and "Work". I will then apply the specified formatting to each identified entity.


Please share the text from the travelogue, and I will begin the process of entity recognition and labeling.





# §§ Experiment ~ ~ NER





Anonymous


 segment\_0001.txt  
Document


 segment\_0002.txt  
Document

 segment\_0003.txt  
Document

 segment\_0004.txt  
Document

 segment\_0005.txt  
Document

 segment\_0006.txt  
Document

 segment\_0007.txt  
Document

Identify and label all entities in the given texts. They are segments of a larger text, in order to stick to the restrictions of this interface. So I will pass segments to you and will ask you to perform the task for each of them. Use the format '##Entity++' for Person, '@@Entity##' for Location, '~~Entity%%' for Date, '+~Entity~~' for Organisation, '%~Entity&&' for Currency, and '\$\$Entity~~' for Work. Take a deep breath and work on this problem step-by-step.

# §§ Experiment ~ ~ NER



## DH Assistant

In this fourth segment of the historical travelogue, I will continue identifying and labeling entities as per your specifications. Here is the labeled text:

1. **Person (##Entity++)**: No specific personal names are mentioned.
2. **Location (@@Entity##)**: There are no specific locations mentioned in this segment.
3. **Organisation ( ++Entity~~)**: The text refers to "Direction" which seems to be an organisational entity, hence it should be labeled as " ++Direction~~".
4. **Date (~~Entity%%)**: There are no specific dates mentioned in this segment.
5. **Currency (%%Entity&&)**: The text mentions "10 Schild - Louisd'or", which should be labeled as "%%10 Schild - Louisd'or&&".
6. **Work (§§Entity~~)**: Specific works mentioned include "Schrebers Säugthiere", "Espers Schmetterlinge", and "Wilhelms Unterhaltungen aus der Naturgeschichte". These should be labeled as "§§Schrebers Säugthiere~~", "§§Espers Schmetterlinge~~", and "§§Wilhelms Unterhaltungen aus der Naturgeschichte~~" respectively.

# §§ Experiment~~NER

(Fortsetzung)

Here is the text with the identified entities labeled:

"Anfangs bat ich zwar die ++Direction~~, mir einige der theuern anzukaufen; allein dies sehr natürliche Begehren fand so viel Anstände, daß ich mich entschloß, für die Hülfsmittel auf eigene Rechnung zu sorgen: man kann leicht ermessen, wie schwer es mir ward, Werke wie §§Schrebers Säugthiere~~, §§Espers Schmetterlinge~~, §§Wilhelms Unterhaltungen aus der Naturgeschichte~~, und eine Menge anderer theurer mathematischen, physikalischen, naturhistorischen Werke zu bezahlen, meine Mineralien-Sammlung zu vermehren, die ich bey dem Unterricht benutzte, und allerley Ausgaben für kleine Nothwendigkeiten zum Anstellen physikalischer Versuche zu bestreiten: denn auch für Mikroskope, Dosen-Sextanten, Meßzirkel, Prismen, geschliffene Gläser, Wagen u.s.w. sorgte ich auf eigene Kosten. Wahrlich hatte ich mein Einkommen und mich der Schule geopfert. Die ++Direction~~ fühlte auch, daß mir einige Entschädigung gebühre, und schenkte mir zweymal %%10 Schild-Louisd'or&&; die Bitte aber, mein Einkommen durch einen regelmäßigen Zuschuß zu mehren, wies sie zurück, obschon sie übrigens nicht viel Bedenken trug, andern Lehrern, die in Schulden gerathen waren, dieselben freygebig zu bezahlen, und dem Rector 200 Franken Besoldung zu schöpfen."

For further analysis or additional segments, please provide the next text.

# §§ Experiment ~ ~ NER



## Anonymous

Now let's save the results of our analysis, and then proceed with further segments. To save our results create a table and fill it with all the entities of segments 1-7. The table should contain the following columns "Entity" the original string you labeled and "Label" the Label (e.g. Person, Date etc.) you assigned.



## DH Assistant

The entities from segments 1-7 have been successfully extracted and compiled into a table. The table contains a total of 32 entries. Here are the first 10 entries for your reference:

Entity	Label
1809	Date
1200 Schweizer-Franken (75 Schildlouisd'or)	Currency
1600 Schweizer Franken (100 Schild-Louisd'or)	Currency
Direction	Organisation
10 Schild → Louisd'or	Currency

# §§ Experiment ~ ~ Funktionsweise?



## **Anonymous**

Thanks a lot, those are all segments. Before we end, I would like to take a look at your inner working. Explain how you proceed in solving the task, step by step and with examples. Do it on the example of the last segment.



## **DH Assistant**

Certainly! Let's break down how I approached the task of identifying and labeling entities in the last segment (Segment 0027) of your historical travelogue. The task involves several steps:

### **Step 1: Reading and Understanding the Text**

- **Action:** I start by reading the entire text segment to understand the context and content.
- **Example from Segment 0027:** The text mentions familial relationships, dates, and personal reflections, indicating a personal and somewhat emotional narrative.

# §§ Experiment ~ ~ Funktionsweise?

(Fortsetzung)

## Step 2: Identifying Entities

- **Action:** I look for specific entities that match the categories you provided: Person, Location, Organisation, Date, Currency, and Work.
- **Example from Segment 0027:**
  - **Person:** I identified "Vater meiner Geliebten" (father of my beloved), "treue Magd" (faithful maid), and "meine edle Freundinn" (my noble friend).
  - **Location:** No specific locations mentioned.
  - **Organisation:** No specific organisations mentioned.
  - **Date:** "18. Febr. 1810" and "19.ten" (19th).
  - **Currency:** No specific currencies mentioned.
  - **Work:** No specific works mentioned.

## Step 3: Applying the Labeling Format

- **Action:** Once the entities are identified, I apply your specified labeling format to each entity.
- **Example from Segment 0027:**
  - **Person:** "##Vater meiner Geliebten++", "##treue Magd++", "##meine edle Freundinn++".

# §§ Experiment ~ ~ Funktionsweise?

(Fortsetzung)

## **Step 4: Reintegrating Labeled Entities into the Text**

- **Action:** I then place these labeled entities back into the text in their original positions.
- **Example from Segment 0027:** The labeled entities are reinserted into the text, replacing the original mentions of people, dates, etc.

## **Step 5: Verification and Adjustment**

- **Action:** I review the text to ensure that all entities are correctly identified and labeled. If necessary, I make adjustments.
- **Example from Segment 0027:** Ensuring that the labels correctly reflect the entities and that no entities were missed.

## **Final Output**

- The final output is the original text with entities labeled according to your format. This process requires careful reading, understanding the context, and accurately applying the labeling format to the identified entities.

# §§ Experiment ~ ~ Normalisieren



## **Anonymous**

now use gpt to normalize the dates in "YYYY-MM-DD" format. if there is no month or no day indicated, leaf the respective position empty in the normalized format, e.g. "1809" should be normalized as "1809", "September 1802" should be normalized as "1802-09". Use your internal "GPT" model to perform the normalization, instead of pattern matching or other programming approaches.



# §§ Experiment ~ ~ Normalisieren



## DH Assistant

To normalize the dates using the GPT model approach, I will interpret each date entity and convert it into the "YYYY-MM-DD" format, leaving the month or day positions empty if they are not specified in the original date entity. Here are the normalized dates:

1. **1809** → "1809"
2. **Winter 1808–1809** → "1808-1809" (Note: The exact months for 'Winter' are not specified, thus only the years are included.)
3. **1805** → "1805"
4. **September 1802** → "1802-09"
5. **9. Juli 1806** → "1806-07-09"
6. **16. August 1806** → "1806-08-16"
7. **1806** → "1806"

# §§ Experiment ~ ~ Normalisieren

(Fortsetzung)

8. **1809** → "1809"
9. **22. Juli 1809** → "1809-07-22"
10. **10. September 1809** → "1809-09-10"
11. **10. Januar 1810** → "1810-01-10"
12. **2. Febr. 1810** → "1810-02-02"
13. **Hornung 1810** → "1810-02" (Note: 'Hornung' is an old term for February.)
14. **14. Februar** → "14. Februar" (Note: The year is not specified.)
15. **18. Febr. 1810** → "1810-02-18"
16. **19. Febr. 1810** → "1810-02-19"
17. **8. März** → "8. März" (Note: The year is not specified.)
18. **Junius** → "Junius" (Note: Neither day nor year is specified.)
19. **May** → "May" (Note: Neither day nor year is specified.)
20. **August** → "August" (Note: Neither day nor year is specified.)
21. **September** → "September" (Note: Neither day nor year is specified.)
22. **14. März 1810** → "1810-03-14"

# §§ Experiment ~ ~ Normalisieren



**Anonymous**

perfect, but from 14. onward you missed some normalizations: "14. Februar" should be transformed to "02-14", out of the context it is fair to assume it is the year 1810 so the normalization should be "1810-02-14". Redo the normalization for all occurrences, which are not normalized in the final form.



**DH Assistant**

Thank you for the clarification. I will redo the normalization for the dates starting from "14. Februar" onwards, assuming the year as 1810 where it's not specified, based on the context you provided. Here are the corrected normalizations:

1. **14. Februar** → "1810-02-14" (Assuming the year 1810)
2. **18. Febr. 1810** → "1810-02-18"
3. **19. Febr. 1810** → "1810-02-19"
4. **8. März** → "1810-03-08" (Assuming the year 1810)
5. **Junius** → "1810-06" (Assuming the year 1810)
6. **May** → "1810-05" (Assuming the year 1810)
7. **August** → "1810-08" (Assuming the year 1810)
8. **September** → "1810-09" (Assuming the year 1810)
9. **14. März 1810** → "1810-03-14"
10. **8. April 1810** → "1810-04-08"
11. **ersten Woche des Julius** → "1810-07" (Assuming the year 1810 and approximating to July)

# §§ Experiment ~ ~ Normalisieren



## **Anonymous**

perfect, let's continue with the entities labeled as "Currency". Therefore identify the two components of currency entities "unit" and "quantity". "Unit" is supposed to be the name of the currency and "quantity" how much of it is described. Enter a dict with "unit" and "quantity" in the column for "Normalization", e.g. "1600 Schweizer Franken" should be normalized to '{"unit": "Schweizer Franken", "quantity": "1600"}'. Take a deep breath and go step by step to be sure you get it right. The "unit" is not always consistent e.g. "Schweizer-Franken" and "Schweizer Franken" are the same unit and should be assigned to the same unit during normalization.

# §§ Experiment ~ ~ Normalisieren



**DH Assistant**

To normalize the currency entities, we'll identify the 'unit' and 'quantity' components within each entity labeled as 'Currency'. I will go step by step to ensure accuracy, considering the variations in unit names. For example, "Schweizer-Franken" and "Schweizer Franken" will be recognized as the same unit.

[...]

Let's proceed with this task.

The normalization of currency entities has been successfully completed. Each currency entity has been split into 'quantity' and 'unit', and variations in unit names have been normalized to ensure consistency. Here's how the updated table looks for a few of the currency entities:

Entity	Label	Normalization
1200 Schweizer-Franken	Currency	{'unit': 'Schweizer Franken', 'quantity': '1200'}
75 Schildlouisd'or	Currency	{'unit': 'Schild-Louisd'or', 'quantity': '75'}
1600 Schweizer Franken	Currency	{'unit': 'Schweizer Franken', 'quantity': '1600'}
100 Schild-Louisd'or	Currency	{'unit': 'Schild-Louisd'or', 'quantity': '100'}
10 Schild-Louisd'or	Currency	{'unit': 'Schild-Louisd'or', 'quantity': '10'}

# §§ Experiment~~

- > Zero-Shot Ansatz (statt Few-Shot Ansatz bei Wang et al)
- > "Nachtrainieren" auf Grundlage der Outputs möglich
- > Verschiedene Entitäten in einem Prompt als Aufgabe formuliert
  - Multiklassen-Problem (statt Einzelklassen bei Wang et al)
- > Versuche zur Normalisierung
- > Verzicht auf Ansatz zur Lösung des Halluzinationsproblems

# ##Evaluation++

> Anhand der Handlabel und gegen vorherige Versuche mit "herkömmlichen" ML-Ansätzen

Modell	Precision	Recall	F1 Score
GPT-NER	0,86	0,29	0,43
Flair (bisher bestes Modell)	0,69	0,11	0,18

# ++Diskussion~~

> Skalierbarkeit, Handhabbarkeit etc. (API?!)

> generativer Character von GPT

GPT hat u.a. eigenständig Schritte zur Normalisierung unternommen: "Den 19.ten" (the 19th, referring to the next day after the 18th of February 1810) is labeled as: ~~-19. Febr. 1810%%



@@Vielen Dank für Ihre  
Aufmerksamkeit##

## Literatur

Sandra Balck, Hermann Beyer-Thoma, Ingo Frank, Anna Ananieva (2023a). "Interlinking Text and Data with Semantic Annotation and Ontology Design Patterns to Analyse Historical Travelogues". In: Digital Humanities Quarterly 17, 3 (Special Issue "Working on and with Categories for Text Analysis: Challenges and Findings from and for Digital Humanities Practices"), 17 pages. <http://www.digitalhumanities.org/dhq/vol/17/3/000726/000726.html> (09.01.2023).

Sandra Balck, Jacob Möhrke, Anna Ananieva (2023b). "Digitale Editionen historischer Reiseberichte: Der Kreislauf historischer Informationen". In: Digital Humanities Day Leipzig 2023 (DHDL 23). Forum für Digital Humanities Leipzig (4. Dezember 2023), Leipzig. <https://doi.org/10.5281/zenodo.10301629>.

Maud Ehrmann, Ahmed Hamdi, Elvys Linhares Pontes, Matteo Romanello, and Antoine Doucet (2023). "Named Entity Recognition and Classification in Historical Documents: A Survey". In: ACM Comput. Surv. 56, 2, Article 27 (February 2024), 47 pages. <https://doi.org/10.1145/3604931>.

Xiaoya Li (2023). "A Practical Survey on Zero-Shot Prompt Design for In-Context Learning". In: Proceedings of the 14th International Conference on Recent Advances in Natural Language Processing (pp. 641-647). INCOMA Ltd., Shoumen, Bulgaria. <https://aclanthology.org/2023.ranlp-1.69>.

Francesco De Toni, Christopher Akiki, Javier De La Rosa, Clementine Fourrier, Enrique Manjavacas, Stefan Schweter, Daniel Van Strien (2022). "Entities, Dates, and Languages: Zero-Shot on Historical Texts with T0." In: Proceedings of BigScience Episode #5 - Workshop on Challenges & Perspectives in Creating Large Language Models (pp. 75-83). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2022.bigscience-1.7>

Shuhe Wang, Xiaofei Sun, Xiaoya Li, Rongbin Ouyang, Fei Wu, Tianwei Zhang, Jiwei Li, and Guoyin Wang (2023). "GPT-NER: Named Entity Recognition via Large Language Models". <https://doi.org/10.48550/arXiv.2304.10428>.

Chengrun Yang, Xuezhi Wang, Yifeng Lu, Hanxiao Liu, Quoc V. Le, Denny Zhou, Xinyun Chen (2023). "Large Language Models as Optimizers". <https://doi.org/10.48550/arXiv.2309.03409>.