

Itinerare erkennen in Reiseberichten. Auszeichnung von Orts- und Personennamen zur Etablierung von Itineraren in Reiseberichten des 19. Jahrhunderts.

Tarjia Alam Nisha, Franziska Pannach, Jörg Wettlaufer
Niedersächsische Akademie der Wissenschaften zu Göttingen,
Georg-August-Universität Göttingen.



Passau, 27.2.2024

DHd 2024 – Passau: Workshop „Generative KI, LLMs und GPT bei digitalen Editionen“



Überblick

- Das Problem
- Das Projekt: www.middle-east-travelers.de (2018 – heute)
- Annotationen als Problemlösung
- LLMs als Annotationstools – Erste Erfahrungen mit Custom GPTs
- Alternativen zu Custom GPTs
- Zusammenfassung / Fazit



Das Problem:

Viele Reiseberichte können als Itinerare gelesen werden, denn Sie nennen Stationen, die der/die Reisende mit eigenen Augen gesehen bzw. besucht hat.

1. Wie kann man diese Orte möglichst ressourcensparend erkennen und identifizieren?
2. Wie kann man diese Orte *automatisiert* von anderen Orten unterscheiden, die im Bericht nur genannt, aber nicht auf der aktuellen Reiseroute liegen?



Das Projekt:

Reiseberichte als literarische Gattung sind u.a. interessant für HistorikerInnen, da es sich (teilweise) um Augenzeugenberichte handelt.

1. Fokus auf Information Retrieval schon edierter bzw. publizierter Texte – Erfassung und Edition bislang unedierter Berichte eher sekundär bzw. nicht geplant.
2. Projekte zu Reiseberichten aus verschiedenen Epochen. Ziel: Bereitstellung von Hilfsmitteln für die Forschung und Erleichterung des Zugangs zu den Texten, Editionen und der Forschungsliteratur.



Dr. Jörg Wettlaufer
NAWG / GCDH / IfDH
Göttingen

GCDH
GÖTTINGEN CENTRE FOR
DIGITAL HUMANITIES



Dr. Deniz Kiliçoğlu
Lichtenberg-Kolleg
Göttingen

Göttingen
Campus



Niedersächsische
Akademie der
Wissenschaften
zu Göttingen



middle-east-travelers.de

[About](#)

[News](#)

[Bibliography](#)

[Travelers](#)

[Itineraries](#)

[Pictures](#)

[Search](#)

[Team](#)

[Privacy](#)





Projekt: www.middle-east-travelers.de

- Erfassung und Bereitstellung von Berichten europäischer Reisender in englischer, deutscher und französischer Sprache in den das Osmanische Reich im langen 19. Jahrhundert (1789 – 1914).
- Kooperationsprojekt mit Deniz Kilinçoğlu (früher Lichtenbergkolleg Göttingen). Antragstellung bei Niedersachsen Vorab/Sprung 2018/19 – aber **keine Förderung**. Seitdem nicht wieder eingereicht... Aber:
 - Publikation einer Webseite mit einer Liste von Reisenden und Sammlung von Primär- und Sekundärliteratur in Zotero.
 - Ca. 800 Texte (auch Übersetzungen) sind vorauss. relevant. Bisher liegen ca. 180 Volltexte aus Google Books und HathiTrust lokal vor.
 - Geplant war: NER über alle Berichte, Erstellung von Itinerarien und Veröffentlichung als Linked Open Data auf der Webseite.

Fragestellung Beispielprojekt: Entstehung von Nationalismus im Osmanischen Reich



Projekt: www.middle-east-travelers.de

Konzept NER: Vermeidung manueller Korrekturen – Akzeptanz von (einigen) Fehlern im OCR und der NER.

Geplante Strategie (2019) bei Förderung des Projekts:

1. Etablierung des Textkorpus
2. Initiale Named Entity Recognition mit Spacy und Nutzung von [Open Semantic Search](#)
3. Nach der Etablierung von Trainingsmaterial, Verbesserung der NER with ML/DL (siehe nächste Folie)
4. Bereitstellung von Facetten und Itinerarien auf der Basis eines kontrollierten Vokabulars (PlaceNames, Personen, Themen)
5. Publikation als Linked Open Data und Verknüpfung mit Wikidata / Factgrid



www.middle-east-travelers.de

Geplante NER Strategie (2019):

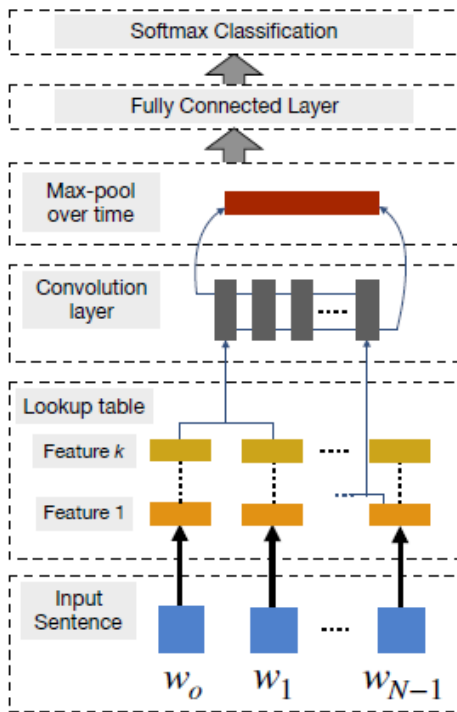


Abbildung 1: Funktionsprinzip eines CNN Frameworks zur Vorhersage der Klasse eines Wortes bei der Erkennung bzw. Zuordnung von Entitäten. Young et al. 2018, Fig. 5 (nach Collobert & Weston 2008).

- Nutzung von Embeddings und einer Kombination aus Convolutional Neural Networks (CNN), Conditional Random Fields (CRF) und Long Short-Term Memory (LSTM) sowie Gazetteers, um Recall und Precision der Erkennung von Entitäten und insbesondere der Ortsnamen zu erhöhen.
- Verknüpfung mit Normdaten und Erweiterung der Open Semantic Search-Plattform für die Veröffentlichung von Linked Open Data.
- Bereitstellung der gewonnenen Daten für interdisziplinäre Forschung über Reiseberichte mit einem Fokus auf Europäische Reisende in den Nahen Osten im 19. Jahrhundert.



Veröffentlichungen:

- Wettlaufer, Jörg & Deniz Kiliñoğlu (2019): Poster: "Linked Open Travel Data: Erschließung gesellschaftspolitischer Veränderungen im Osmanischen Reich im 19. Jahrhundert durch ein multimediales Online-Portal", DHd 2019 Mainz/Frankfurt, Book of Abstracts, S. 315-317.
- Kiliñoğlu, Deniz & Jörg Wettlaufer (2019): Vortrag: "Travels in the 19th-Century Ottoman Empire. A Digital History Project". Gehalten auf dem workshop "Horizons in Digital Research on Religion, Nationalism, and Collective Memory, Forum für interdisziplinäre Religionsforschung (FiReF), Lichtenberg Kolleg Göttingen, 5.12.2019.
- Kiliñoğlu, Deniz & Jörg Wettlaufer (2022): The Reshuffling of Middle Eastern Identities in the Age of Nationalism: Insights from Nineteenth-Century Travelogues, in: On the Way into the Unknown. Comparative Perspectives on the Orient in (Early) Modern Travelogues, ed. by Arno Strohmeyer & Doris Gruber, Studies on Modern Orient, De Gruyter, S. 365-383. doi: <https://doi.org/10.1515/9783110698046>



Das *Problem* der NER Strategie ist weiterhin ungelöst!

„Wie kann man die Orte einer Reiseroute *automatisiert* von anderen Orten unterscheiden, die im Bericht nur genannt, aber nicht auf der aktuellen Route liegen?“

Masterarbeit => Tarjia Alam Nisha



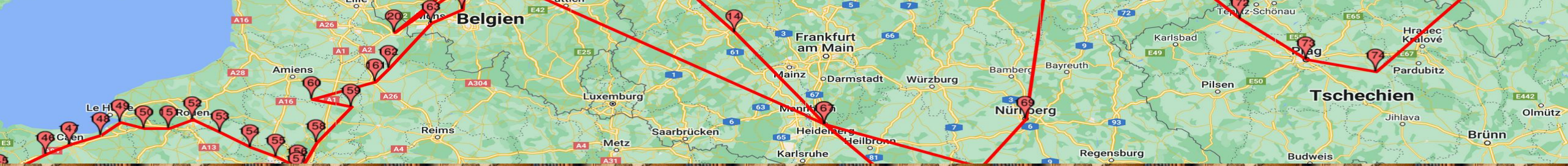


Itinerare erkennen in Reiseberichten.
Auszeichnung von Orts- und Personennamen zur
Etablierung von Itineraren in englischsprachigen
Reiseberichten des 19. Jahrhunderts.

Ein Vorschlag zur automatisierten Auszeichnung von Orten
einer Reisroute mit Hilfe von „Custom GPTs“ von OpenAI

<https://chat.openai.com/g/g-M46uUIBD6-travelogue-annotator> =
<https://s.gwdg.de/xYrXkh>

<https://chat.openai.com/g/g-Uy5ZolqD2-itinerizer> = <https://s.gwdg.de/9kZgMM>



Vorgehensweise:

- Erstellung von Annotationsrichtlinien zur Auszeichnung der Beziehungen zwischen Personen, Orten und Reisebeteiligten unter Berücksichtigung der örtlichen und zeitlichen Beziehung einer Person zu einem gegebenen Ort.
- Auszeichnung eines Beispieltextes anhand der Richtlinien mit Hilfe des Data Analyzers von Chat-GPT 4.0.
- Einrichtung von OpenAI Custom-GPTs mit den Annotationsrichtlinien und (z.T.) Training mit annotierten Beispieltexten (few shot prompting).
- Erweiterung der GPTs um die Fähigkeit, den Geocode der annotierten Orte des Itinerars einzufügen und die Reiseroute auf einer Karte (außerhalb des Chats) anzeigen zu lassen.



Annotationen für die Auszeichnung der Beziehungen zwischen Personen und Orten:

Grounding Characters and Places in Narrative Texts

Sandeep Soni and **Amanpreet Sihra**
University of California, Berkeley
sandeepsoni, amasihra@berkeley.edu

Elizabeth F. Evans
Wayne State University
e.f.evans@wayne.edu

Matthew Wilkens
Cornell University
wilkens@cornell.edu

David Bamman
University of California, Berkeley
dbamman@berkeley.edu

Sandeep Soni, Amanpreet Sihra, Elizabeth Evans, Matthew Wilkens, and David Bamman. 2023. [Grounding Characters and Places in Narrative Text](#). In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 11723–11736, Toronto, Canada. Association for Computational Linguistics.

<https://doi.org/10.48550/arXiv.2305.17561>

Mr. Fogg repaired to [his bedroom]_{TO}, and called out, "Passepartout!"

Last Friday at eight o'clock in the evening, Monsieur Fogg came [home]_{TO} from [his club]_{FROM}

As for Mr. Fogg, he slumbered as peacefully as if he had been in [his serene mansion in [Saville Row]_{NO REL}]_{NO REL}.

On the ninth day after leaving [Yokohama]_{FROM}, Phileas Fogg had traversed exactly one half of the terrestrial globe .

Phileas Fogg wished to go to [Liverpool]_{NO REL} , but the captain would not carry him [there]_{NO REL} .



Recall & Precision bei der Annotationsaufgabe von Soni et al. 2023:

Model	Valid	Rel.	Temp.	Tense
Majority class	79.8	43.5	56.5	71.0
ChatGPT	47.7	30.2	31.4	70.7
GPT-4	83.8	51.0	59.3	69.2
BERT classifier	87.4	56.8	73.3	79.0

Table 3: Accuracy of the model on each task (whether a character/place pair is valid; classifying spatial relationship, temporal span and narrative tense) shown in percentage; The majority class model for every task predicts the most common label in the training set. The performance is reported on the test set.

Category	Precision	Recall	F1
IN	0.68	0.61	0.65
NO REL	0.54	0.68	0.60
TO	0.40	0.53	0.45
FROM	0.43	0.45	0.44
NEAR	0.50	0.24	0.32
THRU	0.20	0.33	0.25

Table 4: The performance of the BERT-based classifier on the spatial prediction task. The table shows the performance broken down per class.



Vorbereitung und Grundlage: Erstellung von Annotationsrichtlinien

Annotation Guidelines

Task objective: The goal is to determine and describe the type of connection or relationship that exists between two specific characters or a specific character (Traveler) and a designated place/ location within a given excerpt. This entails identifying the nature of their association and providing an explanation or description of it. (Which will help us to find valid itineraries of a traveler or group of travelers.)

e.g.: **John Stewart, Esq.** was born in **London** in the year 1749.

e.g.: **I (WILLIAM MACMICHAEL)** shall always reflect with great satisfaction on the adventures of a journey during which I enjoyed the pleasure of the society of my friend **Mr. Legh**.

What is a character? A character refers to a human or a person entity within a narrative that possesses the capability to engage in monologues or dialogues. An automated entity tagger is employed to identify characters, with entities categorized as PER (person) according to the guidelines specified in a particular research paper, being regarded as potential characters.

e.g.: **Mr. Stewart** displayed the most fanciful and yet pleasing taste.

What is a place? A place refers to a defined spatial area that can be physically demarcated. It encompasses locations such as rooms, cities, or specific geographical features. Similar to character identification, candidates for place entities are also recognized automatically using an entity tagger. In this case, entities falling under the categories of LOC (location), FAC (facility), or GPE (geopolitical entity), as outlined in a specific research paper, are regarded as potential places.

e.g.: Mr. Stewart returned to **England** (GPE)

Task details and setup: The excerpts will provide from the source of English travelogues of Ottoman Empire. Within each excerpt, there will be a highlighted section representing a character or characters, as well as a separate highlighted portion representing a place. The identification of these character and place spans will achieve through an automated system, eliminating the need for manual intervention in the process. If an annotator finds a place that was not correctly recognized by the tagger, then annotator will provide the label themselves.

Seven annotations are to be provided by every annotator for any excerpt.

Annotation	Type (Domain)	Possible values
Valid relation	Binary (PER-LOC)	<ol style="list-style-type: none"> TRUE FALSE
Spatial relation	Categorical (PER-LOC)	<ol style="list-style-type: none"> IN NEAR THRU (Through) TO FROM NO REL
Temporal span	Binary (PER-LOC)	<ol style="list-style-type: none"> PUNCTUAL HABITUAL
IsPartOfJourney	Binary (PER-LOC)	<ol style="list-style-type: none"> TRUE FALSE
Valid connection	Binary (PER-PER)	<ol style="list-style-type: none"> TRUE FALSE
Traveling relation/ Travel relationship / Travel connection	Categorical (PER-PER)	<ol style="list-style-type: none"> IN GROUP JOIN GROUP EXIT GROUP NO REL



Durchführung: Annotationschema der Texte

Beispiel für die Annotationen:

**<PER id=0>Mr. Stewart</PER> <Action SpatialRelation=FROM>took leave of</Action> <PLACE
TemporalSpan=HABITUAL, IsPartOfJourney=True>India</PLACE>, and <Action
SpatialRelation=THRU>travelled over</Action> near the whole of <PLACE
TemporalSpan=PUNCTUAL, IsPartOfJourney=True>Persia</PLACE> and <PLACE
TemporalSpan=PUNCTUAL, IsPartOfJourney=True>Turkey</PLACE> on foot; rarely allowing
himself the indulgence of a horse or caravan; and after a variety of accidents and adventures,
<Action SpatialRelation=TO>arrived safely in</Action> his own country. He had fondly cherished
an idea that his doctrines would have been received with enthusiasm by the English; and having
long digested that theory, he was prepared, when <Action SpatialRelation=IN>settled in</Action>
<PLACE TemporalSpan=HABITUAL, IsPartOfJourney=True>London</PLACE>, to make it public.**




1. Schritt GPT: Einrichtung und Konfiguration der GPTs unter ChatGPT 4.0/+

TravelogueAnnotator
Published · Anyone with a link

Unpublished changes Update

Create **Configure**




Name
TravelogueAnnotator

Description
Expert in annotating travelogue excerpts, focusing on character-place relationships.


Instructions
TravelogueAnnotator is specialized in analyzing and annotating connections between characters and places within travelogue excerpts. It focuses on identifying the type of relationship or interaction that exists between a character or characters and a designated place or location. This GPT uses guidelines and descriptions from the user's uploaded documents to precisely annotate travelogues, ensuring accurate and detailed annotations. It aims to support research or study that involves understanding the



Preview



TravelogueAnnotator
Expert in annotating travelogue excerpts, focusing on character-place relationships.


Analyze this travelogue excerpt for chara...
What is the temporal span of the charact...

Can you annotate the spatial relationship...
Determine if this ch 

 Message TravelogueAnnotator... 



2. Schritt GPT: Hinzufügen von Wissen (Annotationsrichtlinien- und Beispiele)

 **TravelogueAnnotator**
Published · Anyone with a link

Unpublished changes ... [Update](#)




[Create](#) [Configure](#)

Conversation starters

- Analyze this travelogue excerpt for character-place relationships. ×
- Can you annotate the spatial relationship in this excerpt? ×
- What is the temporal span of the character's interaction with the place in this text? ×
- Determine if this character's mention of a place is part of their journey. ×
- ×


Knowledge

If you upload files under Knowledge, conversations with your GPT may include file contents. Files can be downloaded when Code Interpreter is enabled

-  **Annotation_description_Fi...**
PDF
-  **Sample_Eamples.pdf**
PDF
-  **AnnotationGuidelines-Fin...**
PDF

[Upload files](#)


Preview



TravelogueAnnotator

Expert in annotating travelogue excerpts, focusing on character-place relationships.


- Analyze this travelogue excerpt for character-...
- What is the temporal span of the character's i...
- Can you annotate the spatial relationship in th...
- Determine if this character's mention of a plac...

 Message TravelogueAnnotator... ↑

?




3. Schritt GPT: Konfiguration der Fähigkeiten und zusätzliche Einstellungen


 TravelogueAnnotator
Published · [Anyone with a link](#)


Unpublished changes ... [Update](#)

[Create](#) [Configure](#)

downloaded when Code Interpreter is enabled

 Annotation_description_Fi... PDF

 Sample_Eamples.pdf PDF

 AnnotationGuidelines-Fin... PDF

[Upload files](#)

Capabilities

- Web Browsing
- DALL-E Image Generation
- Code Interpreter [?](#)


Actions

[Create new action](#)

Additional Settings

- Use conversation data in your GPT to improve our models

Preview



TravelogueAnnotator

Expert in annotating travelogue excerpts, focusing on character-place relationships.

Analyze this travelogue excerpt for character-...

What is the temporal span of the character's i...

Can you annotate the spatial relationship in th...

Determine if this character's mention of a plac...

[Message TravelogueAnnotator...](#) [↑](#) [?](#)



4. Schritt GPT: Trainieren des GPTs mit annotierten Beispieldaten (finetuning)

Anonymous

Check my result. you miss the <Action> part before a place entity and this one is not related to the actual journey<Action SpatialRelation=TO>travelling</Action>(describing his wish). He determined on acquiring the languages of the <PLACE TemporalSpan=HABITUAL, IsPartOfJourney=True>country</PLACE>, and to become an interpreter, until by the exercise of economy he had saved £3000, then to purchase an annuity and <Action>return to</Action> <PLACE TemporalSpan=HABITUAL, IsPartOfJourney=True>England</PLACE>, before the climate had injured his health. The intention upon which this resolution was grounded, arose from a plan he formed, when released from actual dependence, of travelling over the whole world, to trace the causes of human misery, and the causes of human welfare, and of determining how far the latter might be brought to perfection

TravelogueAnnotator

Thank you for pointing out the missed <Action> tag and the clarification regarding the non-journey-related action of traveling. Based on your feedback, here's the revised annotation:

Revised Annotated Data:

At the age of thirteen <PER id=0>he</PER> was <Action SpatialRelation=TO>removed</Action> to the <PLACE TemporalSpan=HABITUAL, IsPartOfJourney=False>CharterHouse</PLACE>; even there his aversion to books continued; he neglected every opportunity for study, at times only composing moral essays or themes, which, as the master declared, elicited genius superior to that of any boy in the class. Still the young scholar was pronounced a dunce; and his <PER id=1, ValidConnection=True, TravelConnection=False>father</PER> was advised to <Action SpatialRelation=TO>send</Action> him abroad, in the hopes of removing his indolent habits. This was accordingly done, and the youth <Action SpatialRelation=TO>sailed</Action> to the <PLACE TemporalSpan=HABITUAL, IsPartOfJourney=True>East Indies</PLACE> as a writer in the Company's service. It was in that <PLACE TemporalSpan=HABITUAL, IsPartOfJourney=True>country</PLACE> where his character and energies began to unfold themselves: his previous education not having left any heavy over-burdened impressions on his mind; his faculties were free and unfixed. Thought and reflection now enabled <PER id=0>Mr. Stewart</PER> to exercise his



Beispieltexte (in Auszügen):

1. Brande, William Thomas. *The Life and Adventures of the Celebrated Walking Stewart: Including His Travels in the East Indies, Turkey, Germany, & America*. E. Wheatley, 1822.
2. Barker, Henry Aston, Burford, J., Pomardi, Simone, and Panorama. *Description of the View of Athens and Surrounding Country: Now Exhibiting in Henry Aston Barker and John Burford's Panorama, Strand; with an Improved Explanation, Giving a Complete Outline of the Whole Picture, with Numbers and References*. H. Barker, 1818.
3. Chandler, Richard. *Travels in Greece: Or an account of a tour made at the expense of the Society of Dilettanti*, 1817.
4. G. A. Oliviers *Reisen durch das Türkische Reich, Ägypten und Persien...* in: Matthias Christian Sprengel. *Bibliothek der neuesten und wichtigsten Reisebeschreibungen zur Erweiterung der Erdkunde nach einem systematischen Plane bearbeitet*, Band 6, Verlag des Industrie-Comptoirs, 1802. [Voyage dans l'Empire Othoman, l'Egypte et Perse fait par Ordre du Gouvernement pendant six premieres années de la Republique, par G.A. Olivier, 2 Vol.]
5. Radzikowski, Piotr (Hg.), *Reyße-Beschreibung Niclas von Popplau, Ritters, bürtig von Breßlau, Übers. von Vorw., Bemerkungen und Annex nach der poln. Ausg.: Piotr Radzikowski, Danuta Radzikowska, Kraków 1998 (Prace Instytutu Historii Wyższej Szkoły Pedagogicznej w Kielcach; 8).*



Ergebnis eines Annotationsexperiments mit Hilfe des TravelogueAnnotators (Version Januar 2024):

TravelogueAnnotator

Original Data:

"Voyage to the strait of Gibraltar—Custom of the sailors—Our passage through the strait—A species of porpoise described. The sun-set remarkable—Ancient accounts of it—The cause. We embarked at Gravesend on the 9th of June, 1764, in the Anglicana, a ship carrying sixteen guns, and thirty-two men, burden about three hundred tons. The commander was Capt. John Stewart; the price of our passage to Turkey sixty guineas. We had a fair wind; but our pilot, being in liquor, did not sail that evening. On Whitsunday, early in the morning, we got under way with a brisk gale, and arrived in the Downs about four in the afternoon. The next day we weighed anchor again, and proceeded to Falmouth to complete our cargo. We were detained there from the 17th to the 24th, when we recovered our anchor with some difficulty, and got clear of the harbour. A signal was made for a pilot, but he did not come on board soon enough to be of use. The wind had been very high while we were in the port of Falmouth, and the weather was still unsettled."

<Author> Richard Chandler </Author>

<Title>

TRAVELS IN ASIA MINOR, AND GREECE: OR. An Account of a Tour MADE AT THE EXPENSE OF THE SOCIETY OF DILETTANTI. By R. CHANDLER, D.D. FELLOW OF MAGDALEN COLLEGE; AND OF THE SOCIETY OF ANTIQUARIES.

</Title> <Published>1817</Published>

Annotated Data:

"Voyage to the <PLACE TemporalSpan=HABITUAL, IsPartOfJourney=TRUE, Lat=35.9912, Lon=-5.6103>strait of Gibraltar</PLACE>—Custom of the sailors—Our <Action SpatialRelation=THROUGH>passage through</Action> the strait—A species of porpoise described. The <Action SpatialRelation=TO>sun-set</Action> remarkable—Ancient accounts of it—The cause. <PER id=0>We</PER> <Action SpatialRelation=IN>embarked</Action> at <PLACE TemporalSpan=PUNCTUAL, IsPartOfJourney=TRUE, Lat=51.4414, Lon=0.3680>Gravesend</PLACE> on the 9th of June, 1764, in the Anglicana, a ship carrying sixteen guns, and thirty-two men, burden about three hundred tons. The commander was <PER id=2>Capt. John Stewart</PER>; the price of <PER id=0>our</PER> passage to <PLACE TemporalSpan=PUNCTUAL, IsPartOfJourney=TRUE, Lat=39.9334, Lon=32.8597>Turkey</PLACE> sixty guineas. <PER id=0>We</PER> had a fair wind; but <PER id=3>our pilot</PER>, being in liquor, did not <Action SpatialRelation=FROM>sail</Action> that evening. On Whitsunday, early in the morning, <PER id=0>we</PER> got under way with a brisk gale, and <Action SpatialRelation=IN>arrived</Action> in <PLACE TemporalSpan=PUNCTUAL, IsPartOfJourney=TRUE, Lat=51.1297, Lon=1.3134>the Downs</PLACE> about four in the afternoon. The next day <PER id=0>we</PER> weighed anchor again, and <Action SpatialRelation=TO>proceeded</Action> to <PLACE TemporalSpan=HABITUAL, IsPartOfJourney=TRUE, Lat=50.1544, Lon=-5.0711>Falmouth</PLACE> to complete <PER id=0>our</PER> cargo. <PER id=0>We</PER> were detained there from the 17th to the 24th, when <PER id=0>we</PER> <Action SpatialRelation=FROM>recovered</Action> <PER id=0>our</PER> anchor with some difficulty, and got clear of the harbour. A signal was made for a pilot, but he did not come on board soon enough to be of use. The wind had been very high while <PER id=0>we</PER> were in the port of <PLACE TemporalSpan=HABITUAL, IsPartOfJourney=TRUE, Lat=50.1544, Lon=-5.0711>Falmouth</PLACE>, and the weather was still unsettled."



Blau: Itinerar von ChatGPT
Rot: Korrekte Reiseroute



Ergebnisse des ersten Experiments:

- Korrekte Erkennung der Orte und weitgehend korrekte Auszeichnung der Qualität eines Ortes als Teil der Reise oder als sonstige Ortserwähnung, aber: Gibraltar als Reiseziel wird in die Mitte der Reiseroute gesetzt, nicht an das Ende. Gründe dafür müssen wir noch eruieren...
- „Geocoding“ funktioniert für normale Ortsnamen einwandfrei. Es bleibt zu testen, wie zuverlässig historische Ortsbezeichnungen und/oder Gebäude bzw. Orte innerhalb von (größeren) Orten erkannt werden.



Ergebnis einer Annotationsaufgabe mit einer anderen Sprache / Sprachstufe:

Reysse-Beschreibung Niclas von Popplau, Ritters, bürtig von Breßlau Am Abend purificationis oder Mariae Lichtmeß im 1483. Jahre bin ich von Wienn aus Kayl. Maytt. Hof gezogen mit 5 Roßen und so viel Gesinde, zusampt dem Bischoff von Gran und Förlin (: welche mit mir zugleich auszogen:), gen Melck, eine Stadt in Österreich kommen. ...

You
Analyze this travelogue excerpt for character-place relationships.

TravelogueAnnotator
It appears that the document does not contain a specific travelogue excerpt for analysis. Could you please provide the excerpt you would like analyzed, or specify if there is another document containing the excerpt?



Itinerizer
Zur Annotation des bereitgestellten Textauszugs aus der "Reysse-Beschreibung Niclas von Popplau" verwende ich die im Annotationsschema festgelegten Richtlinien. Ich werde den Text Schritt für Schritt analysieren und annotieren, wobei ich die Beziehungen zwischen den Charakteren und den Orten identifiziere. Anschließend erstelle ich eine Karte im HTML-Format, die die Reiseroute von Niclas von Popplau darstellt.



Zuerst identifiziere ich die Orte und Personen im Text und weise ihnen entsprechende Tags und Attribute zu. Anschließend werde ich die notwendigen Schritte zur Erstellung der Karte durchführen. Ich beginne jetzt mit der Annotation und melde mich zurück, sobald diese abgeschlossen ist. Die Annotation des Textauszugs aus der "Reysse-Beschreibung Niclas von Popplau" ist nun abgeschlossen. Hier ist der annotierte Text:

```
php Copy code  
  
<PER id=0>Niclas von Popplau</PER>, Ritters, bürtig von Breßlau. Am Abend purificationis
```

Zusätzlich habe ich eine Karte im HTML-Format erstellt, die die Orte darstellt, die Niclas von Popplau auf seiner Reise besucht hat. Sie können die Karte über den folgenden Link herunterladen:

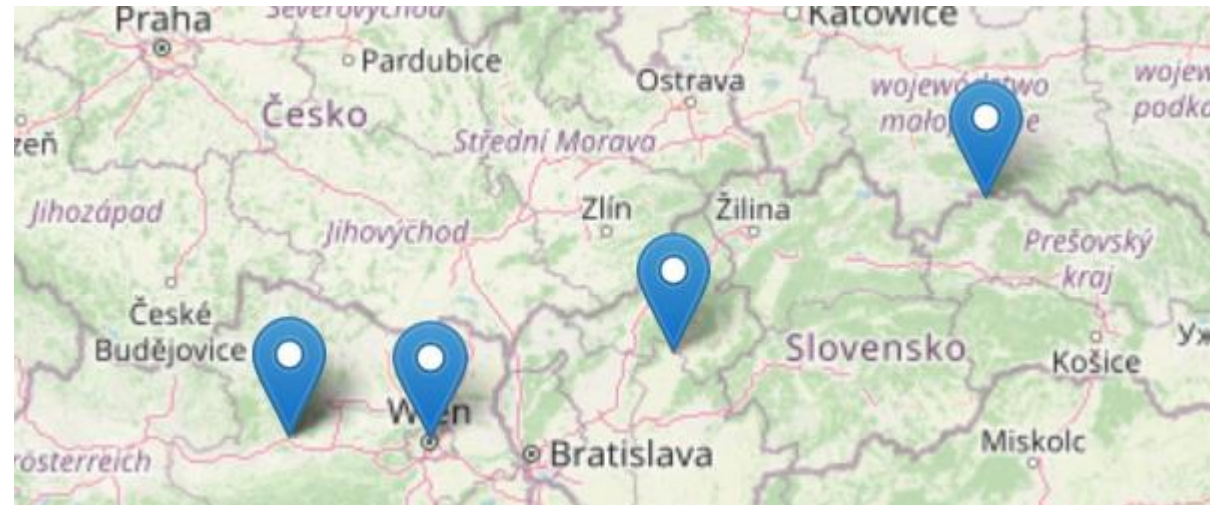
[Download der Reisekarte von Niclas von Popplau \[-\]](#)





Ergebnis einer Annotationsaufgabe mit einer anderen Sprache / Sprachstufe:

Reysse-Beschreibung Niclas von Popplau, Ritters, bürtig von Breßlau Am Abend purificationis oder Mariae Lichtmeß im 1483. Jahre bin ich von Wienn aus Kayl. Maytt. Hof gezogen mit 5 Roßen und so viel Gesinde, zusamt dem Bischoff von Gran und Förlin (: welche mit mir zugleich auszogen:), gen Melck, eine Stadt in Österreich kommen. ...



2/3

1

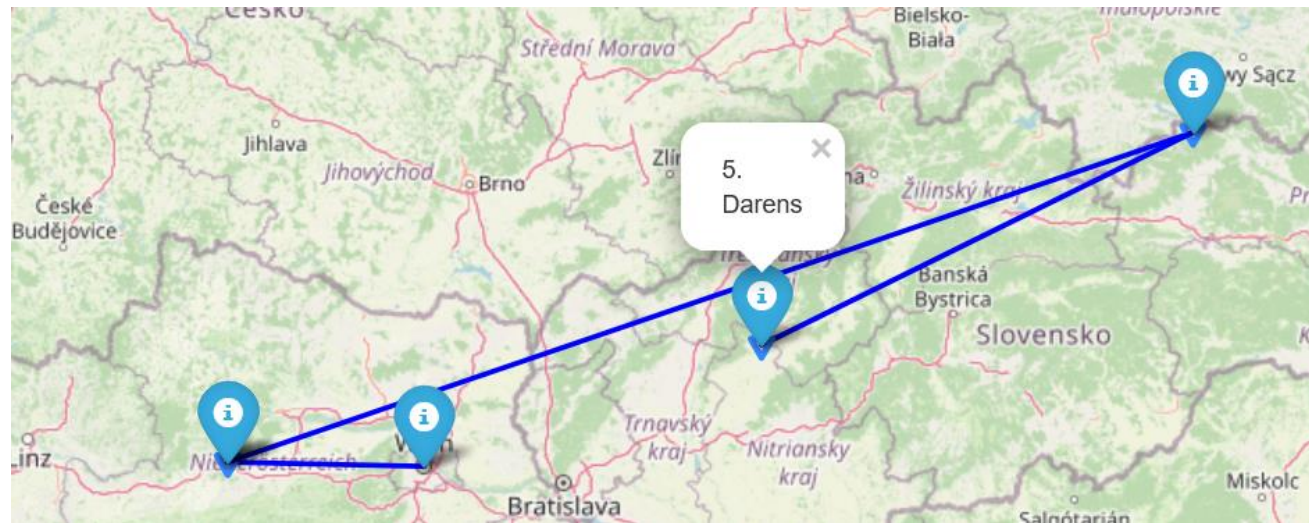
5

4



Ergebnis einer Annotationsaufgabe mit einer anderen Sprache / Sprachstufe:

Prompt: Erstelle die Karte erneut und verbinde in der neuen Version die Orte durch Linien, die einen Pfeil enthalten, der dem Verlauf der Reise entspricht. Nummeriere auch die Datenpunkte, um den Verlauf der Reise zu verdeutlichen. Beginne mit 1 für den zuerst genannten Ort, der auf der Reiseroute liegt und zähle dann hoch bis zum letzten Ort.



2/3

1

5

4



G. N. Olivier
 Reise
 durch das
 Türkische Reich,
 Egypten und Persien,
 während der ersten sechs Jahre der französischen
 Republik
 oder
 von 1792 bis 1798.
 herausgegeben
 von
 M. C. Sprengel.
 Erster Theil.
 Weimar,
 im Verlage des Landes-Industrie-Comptoirs,
 1802.
 Bibliothek
 der
 neuesten und wichtigsten
 Reisebeschreibungen
 zur
 Erweiterung der Erdkunde
 nach einem
 systematischen Plane bearbeitet,
 und in Verbindung
 mit einigen andern Quellen gesammelt
 und
 herausgegeben
 von
 M. C. Sprengel.
 Sechster Band.
 Weimar,
 im Verlage des Landes-Industrie-Comptoirs,
 1802.

308 Abreise von Constantinopel
 den oben erwähnten Theilen des türkischen Reichs zu
 gelangen.

Der Bürger Descorches erleichterte unsere Absicht,
 so viel es ihm möglich war, und wir verließen Constans
 tinopel, gegen Ende des Novembers im Jahr 2, auf eis
 nem Ragusanischen Schiffe, welches ein französischer
 Kaufmann befrachtet hatte. Der Capitain sollte in dem
 Meerbusen von Mundania ankeren, um einige Ballen
 Seide von Brussa einzunehmen; von da sollte er nach
 Smyrna fahren, und uns bey den Dardanellen absetzen.

Der Tag war schön, und das Meer vollkommen
 ruhig, und der Capitain benutzte den ersten Nordwind,
 um von Galata abzusegeln, und wir legten des Abends
 bey dem Dorfe Prinkipos bey. Da wir uns der Winds
 stille wegen, zwey Tage aufhalten mußten, beschäftigten
 wir uns damit, Taucher, Schnepfen und Rebhühner
 zu schießen, und Gräser und andre Pflanzen zu samm
 meln. Den 9ten kamen wir eine halbe Meile von dem
 Dorfe Ghemlek vor Anker.

Das Dorf Ghemlek, welches die Griechen Kios
 nennen, liegt nördlich am Ende des Meerbusens. Es
 ist auf den Ruinen des alten Cius erbaut, und hat
 etwa 2000 Einwohner; sie bestehen größtentheils aus
 Griechen, und beschäftigen sich mit Ackerbau, und eis
 nige wenige mit der Schifffahrt. Einige müssen auf den
 türkischen Schiffswerften arbeiten. Bey unserer Ans
 kunft war ein Fahrgew von 24 Kanonen, unter der

Matthias Christian Sprengel. Bibliothek
 der neuesten und wichtigsten
 Reisebeschreibungen zur Erweiterung
 der Erdkunde nach einem
 systematischen Plane bearbeitet, Band 6
 Verlag des Industrie-Comptoirs, 1802

Prompt: Annotiere den folgenden Ausschnitt aus deinem Reisebericht nach dem vorgegebenen Annotationsschema und stelle eine Karte im html-Format zum Download zur Verfügung: "Der Bürger Descorches erleichterte unsere Absicht, so viel es ihm möglich war, und wir verließen **Konstantinopel**, gegen Ende des Novembers im Jahr 2, auf einem Ragusanischen Schiffe, welches ein französischer Kaufmann befrachtet hatte. Der Kapitän sollte in dem Meerbusen von **Mundanla** anlegen, um einige Ballen Seide von **Brussa** einzunehmen; von da sollte er nach **Smyrna** fahren, und bei den **Dardanellen** ablegen. Der Tag war schön, und das Meer vollkommen ruhig, und der Kapitän benutzte den ersten Nordwind, um von **Galata** abzusegeln, und wir legten des Abends bei dem Dorfe-**Prinkipos** bei. Da wir uns der Windstille wegen, zwei Tage aufhalten mussten, beschäftigten wir uns damit, Taucher, Schnepfen und Rebhühner zu schießen, und Gräser und andere Pflanzen zu sammeln. Den 9ten kamen wir eine halbe Meile von dem Dorfe **Gemlik** vor Anker. Das Dorf **Gemlik**, welches die Griechen **Kios** nennen, liegt nördlich am Ende des Meerbusens. Es ist auf den Ruinen des alten **Cius** erbaut, und hat etwa 2000 Einwohner; sie bestehen größtenteils aus Griechen, und beschäftigen sich mit Ackerbau, und einige wenige mit der Schifffahrt. Einige müssen auf den türkischen Schiffswerften arbeiten."



Ergebnisse:

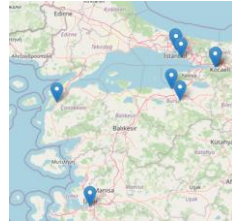
Prompt: Annotiere den folgenden Ausschnitt aus diesem Reisebericht nach dem vorgegebenen Annotationschema und stelle eine Karte im html-Format zum Download zur Verfügung:

Itinerizer


Die Annotation des Reisebericht-Ausschnittes wurde erfolgreich durchgeführt. Sie können die Karte, die die Reise visualisiert, über den folgenden Link herunterladen:

[Download der Reisekarte im HTML-Format \[-\]](#)

< 3 / 3 >   




Itinerizer

 Error analyzing ▾


Es scheint, dass ich derzeit keine Koordinaten für die genannten Orte abrufen kann. Ich werde stattdessen meine eigenen Wissensquellen verwenden, um die Koordinaten zu bestimmen:

- **Konstantinopel (Istanbul):** Breitengrad 41.0082, Längengrad 28.9784
- **Brussa (Bursa):** Breitengrad 40.1826, Längengrad 29.0669
- **Smyrna (Izmir):** Breitengrad 38.4192, Längengrad 27.1287
- **Dardanellen:** Breitengrad 40.0379, Längengrad 26.4086
- **Galata (Teil von Istanbul):** Breitengrad 41.0214, Längengrad 28.9744
- **Dorf Prinkipos (Büyükdada):** Breitengrad 40.8763, Längengrad 29.0916
- **Gemlik:** Breitengrad 40.4309, Längengrad 29.1597

Nun werde ich eine Karte im HTML-Format erstellen, die die genannten Orte und ihre Verbindungen zeigt.

 Error analyzing ▾

Leider gab es ein Problem beim Erstellen der Karte. Ich werde einen weiteren Versuch unternehmen, um die Karte korrekt zu erstellen und als HTML-Datei zur Verfügung zu stellen. Bitte einen Moment Geduld. [-]

 Error analyzing ▾

Es tut mir leid, aber es scheint, dass ich weiterhin technische Schwierigkeiten habe, die Karte zu erstellen. Leider kann ich derzeit keine Karte im HTML-Format zur Verfügung stellen. Falls Sie weitere Anfragen oder einen anderen Text zur Annotation haben, stehe ich Ihnen gerne zur Verfügung. [-]

< 4 / 4 >   

Annotierter Text:

```
<PER id=0>Der Bürger Descorches</PER> <Action SpatialRelation=FROM>erleichterte unsere Absicht, so viel es ihm möglich war, und wir verließen</Action> <PLACE TemporalSpan=PUNCTUAL, IsPartOfJourney=TRUE, lat=41.0082, lon=28.9784>Konstantinopel</PLACE>, gegen Ende des Novembers im Jahr 2, auf einem Ragusanischen Schiffe, welches ein französischer Kaufmann befrachtet hatte. Der Kapitän sollte in dem Meerbusen von Mundania anlegen, um einige Ballen Seide <Action SpatialRelation=FROM>von</Action> <PLACE TemporalSpan=PUNCTUAL, IsPartOfJourney=TRUE, lat=40.1826, lon=29.0669>Brussa</PLACE> einzunehmen; <Action SpatialRelation=TO>von da sollte er nach</Action> <PLACE TemporalSpan=PUNCTUAL, IsPartOfJourney=TRUE, lat=38.4192, lon=27.1287>Smyrna</PLACE> fahren, und <Action SpatialRelation=NEAR>bei</Action> den <PLACE TemporalSpan=PUNCTUAL, IsPartOfJourney=TRUE, lat=40.0379, lon=26.4086>Dardanellen</PLACE> ablegen. Der Tag war schön, und das Meer vollkommen ruhig, und der Kapitän benutzte den ersten Nordwind, um <Action SpatialRelation=FROM>von</Action> <PLACE TemporalSpan=PUNCTUAL, IsPartOfJourney=TRUE, lat=41.0214, lon=28.9744>Galata</PLACE> abzusegeln, und <Action SpatialRelation=IN>wir legten des Abends bei</Action> dem <PLACE TemporalSpan=PUNCTUAL, IsPartOfJourney=TRUE, lat=40.8763, lon=29.0916>Dorf Prinkipos</PLACE> bei. Da wir uns der Windstille wegen, zwei Tage aufhalten mussten, beschäftigten wir uns damit, Taucher, Schnepfen und Rebhühner zu schießen, und Gräser und andere Pflanzen zu sammeln. Den 9ten kamen <Action SpatialRelation=NEAR>wir eine halbe Meile von dem Dorfe</Action> <PLACE TemporalSpan=PUNCTUAL, IsPartOfJourney=TRUE, lat=40.4309, lon=29.1597>Gemlik</PLACE> vor Anker. <PLACE TemporalSpan=PUNCTUAL, IsPartOfJourney=TRUE, lat=40.4309, lon=29.1597>Das Dorf Gemlik</PLACE>, welches die Griechen Kios nennen, liegt nördlich am Ende des Meerbusens. Es ist auf den Ruinen des alten Cius erbaut, und hat etwa 2000 Einwohner; sie bestehen größtenteils aus Griechen, und beschäftigen sich mit Ackerbau, und einige wenige mit der Schifffahrt. Einige müssen auf den türkischen Schiffswerten arbeiten.*
```



Ergebnisse:

Annotiere den folgenden Ausschnitt aus diesem Reisebericht nach dem vorgegebenen Annotationschema und stelle eine Karte im html-Format zum Download zur Verfügung:



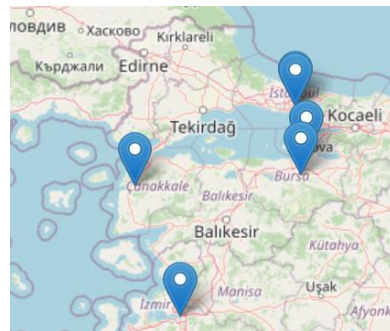
TravelogueAnnotator

Die Annotation des bereitgestellten Reiseberichts wurde gemäß dem vorgegebenen Schema durchgeführt. Hier ist das annotierte Ergebnis:

```
php Copy code  
<PER id=0>Bürger Descorches</PER> erleichterte <PER id=1>unsere</PER> Absicht, so viel
```

Ich habe auch eine Karte erstellt, die den Reiseweg von Konstantinopel bis Gemlik, einschließlich der Zwischenstopps, visualisiert. Die Karte ist im HTML-Format gespeichert und kann über den folgenden Link heruntergeladen werden:

[Download der Reisekarte](#)



<PER id=0>Bürger Descorches</PER> erleichterte <PER id=1>unsere</PER> Absicht, so viel es ihm möglich war, und <PER id=1>wir</PER> verließen <PLACE TemporalSpan=HABITUAL, IsPartOfJourney=TRUE, lat=41.0082, lon=28.9784>Konstantinopel</PLACE>, gegen Ende des Novembers im Jahr 2, auf einem Ragusanischen Schiffe, welches ein <PER id=2>französischer Kaufmann</PER> befrachtet hatte. **Der <PER id=3>Kapitän</PER> sollte in dem <PLACE TemporalSpan=PUNCTUAL, IsPartOfJourney=TRUE>Meerbusen von Mundanla</PLACE> anlegen, um einige Ballen Seide von <PLACE TemporalSpan=PUNCTUAL, IsPartOfJourney=TRUE, lat=40.1828, lon=29.0665>Brussa</PLACE> einzunehmen; von da sollte er nach <PLACE TemporalSpan=PUNCTUAL, IsPartOfJourney=TRUE, lat=38.4192, lon=27.1287>Smyrna</PLACE> fahren, und bei den <PLACE TemporalSpan=PUNCTUAL, IsPartOfJourney=TRUE, lat=40.0379, lon=26.4086>Dardanellen</PLACE> ablegen. Der Tag war schön, und das Meer vollkommen ruhig, und der <PER id=3>Kapitän</PER> benutzte den ersten Nordwind, um von <PLACE TemporalSpan=HABITUAL, IsPartOfJourney=TRUE, lat=41.0247, lon=28.9744>Galata</PLACE> abzusegeln, und <PER id=1>wir</PER> legten des Abends bei dem <PLACE TemporalSpan=PUNCTUAL, IsPartOfJourney=TRUE>Dorfe-Prinkipos</PLACE> bei. Da <PER id=1>wir</PER> uns der Windstille wegen, zwei Tage aufhalten mussten, beschäftigten wir uns damit, Taucher, Schnepfen und Rebhühner zu schießen, und Gräser und andere Pflanzen zu sammeln. Den 9ten kamen <PER id=1>wir</PER> eine halbe Meile von dem <PLACE TemporalSpan=PUNCTUAL, IsPartOfJourney=TRUE, lat=40.4264, lon=29.1597>Dorfe Gemlik (Kios)</PLACE> vor Anker.**



Ergebnisse:

Prompt: „Erstelle die Karte erneut und verbinde in der neuen Version die Orte durch Linien, die einen Pfeil enthalten, der dem Verlauf der Reise entspricht. Nummeriere auch die Datenpunkte, um den Verlauf der Reise zu verdeutlichen. Beginne mit 1 für den zuerst genannten Ort, der auf der Reiseroute liegt und zähle dann hoch bis zum letzten Ort.“

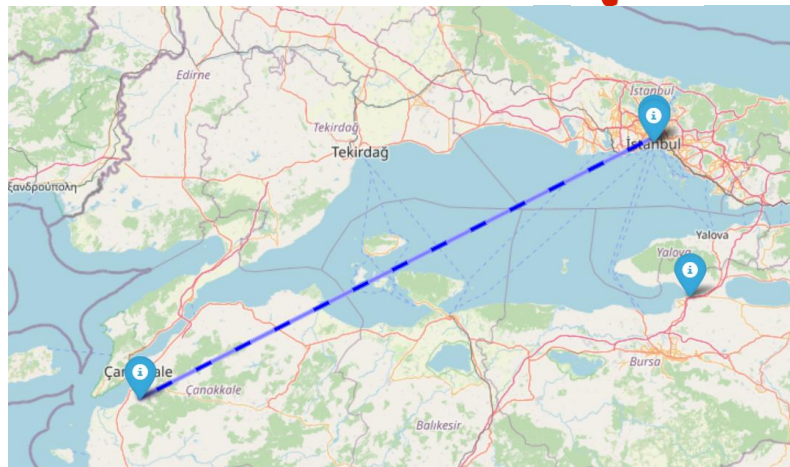
TravelogueAnnotator

Die Karte wurde aktualisiert, um die Reiseroute mit Linien und Pfeilen darzustellen, die den Verlauf der Reise anzeigen. Die Orte sind nummeriert, beginnend mit 1 für den zuerst genannten Ort auf der Reiseroute, und steigen dann entsprechend der Reihenfolge der Reise an.

Die aktualisierte Karte können Sie über den folgenden Link herunterladen:

[Download der aktualisierten Reisekarte \[-\]](#)

🗖 🗑 🔄

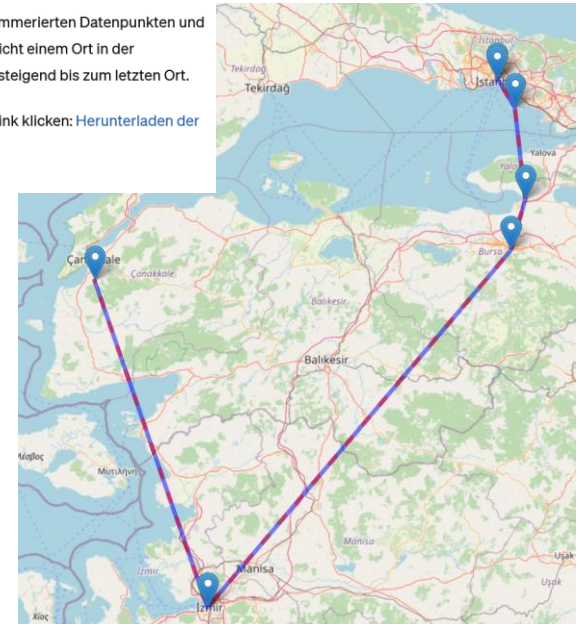
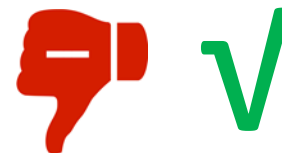


Itinerizer

Die Karte wurde erfolgreich erstellt und zeigt die Reiseroute mit nummerierten Datenpunkten und einer Linie, die die Orte verbindet. Jeder Punkt auf der Karte entspricht einem Ort in der Reihenfolge der Reise, beginnend mit "1" für den ersten Ort und aufsteigend bis zum letzten Ort.

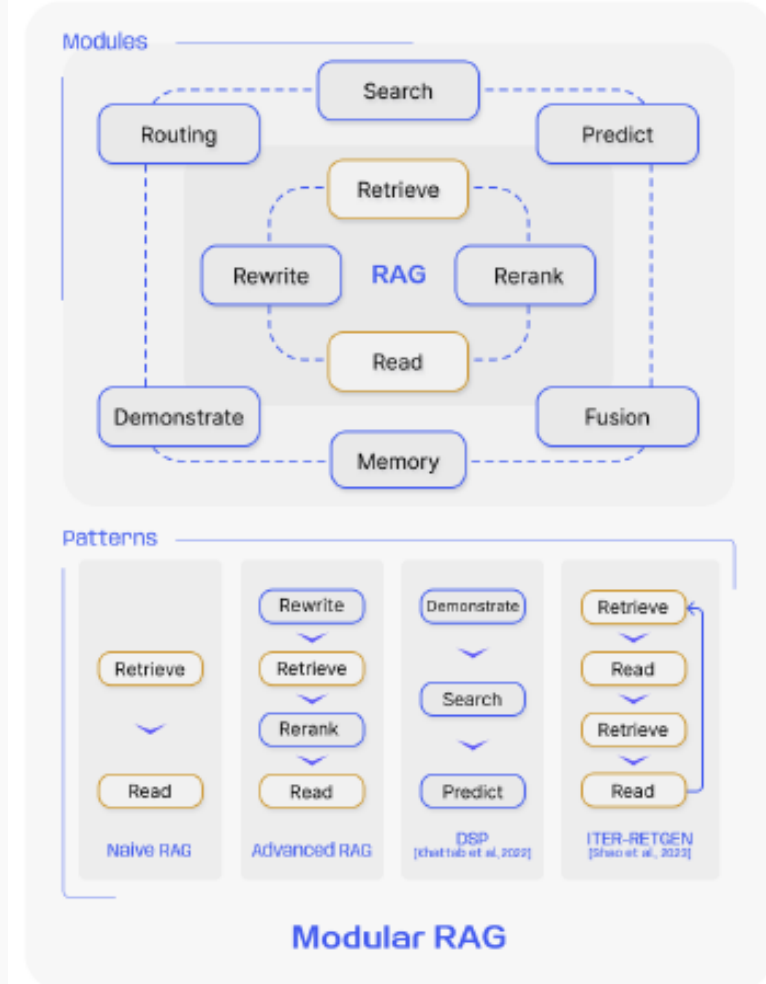
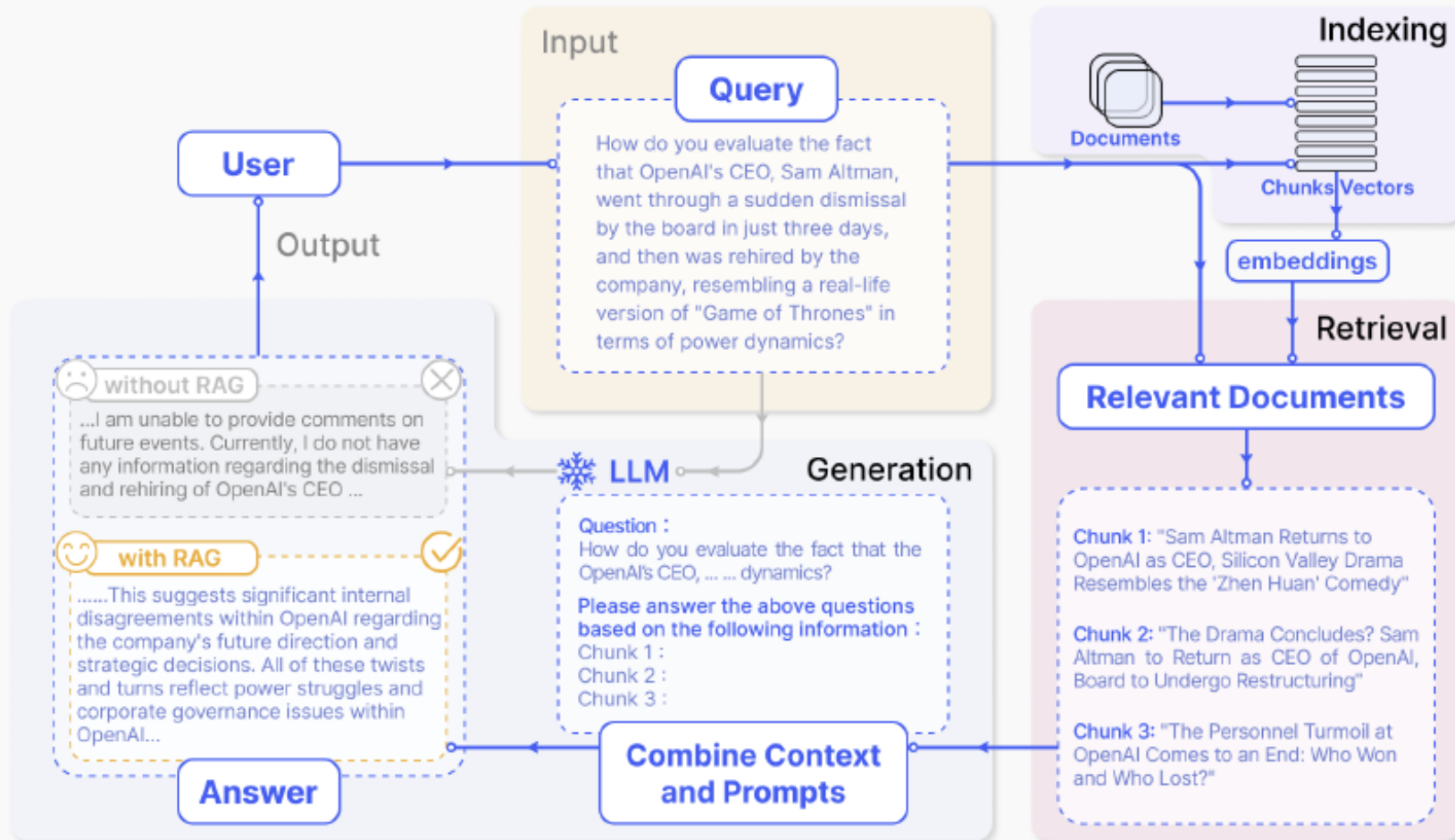
Sie können die Karte herunterladen, indem Sie auf den folgenden Link klicken: [Herunterladen der annotierten Karte. \[-\]](#)

🗖 🗑 🔄





Alternativen? Retrieval Augmented Generation





Alternativen? Nutzung von APIs und freier LLMs (llama2 etc.)...

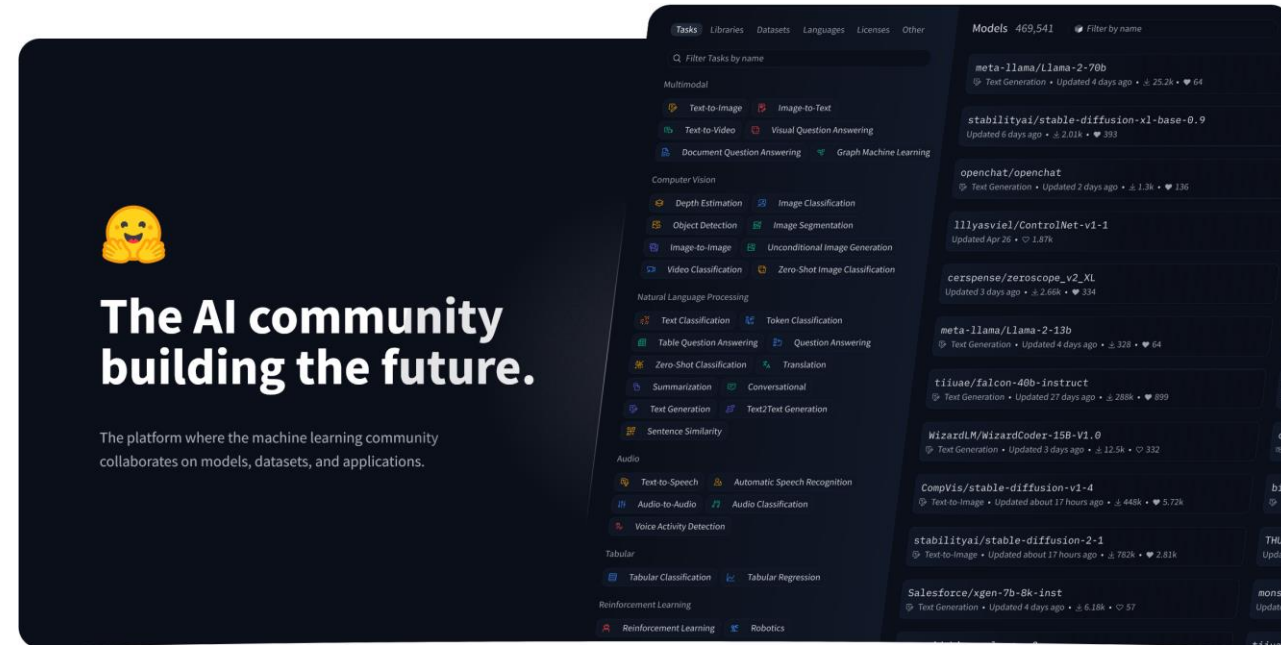
<https://huggingface.co/>

Open Source LLMs and more...

e.g.

- Llama2 70B Chatbot (Meta)
- ...

Siehe: Bager & Gieselmann 2023, ct 23 oder





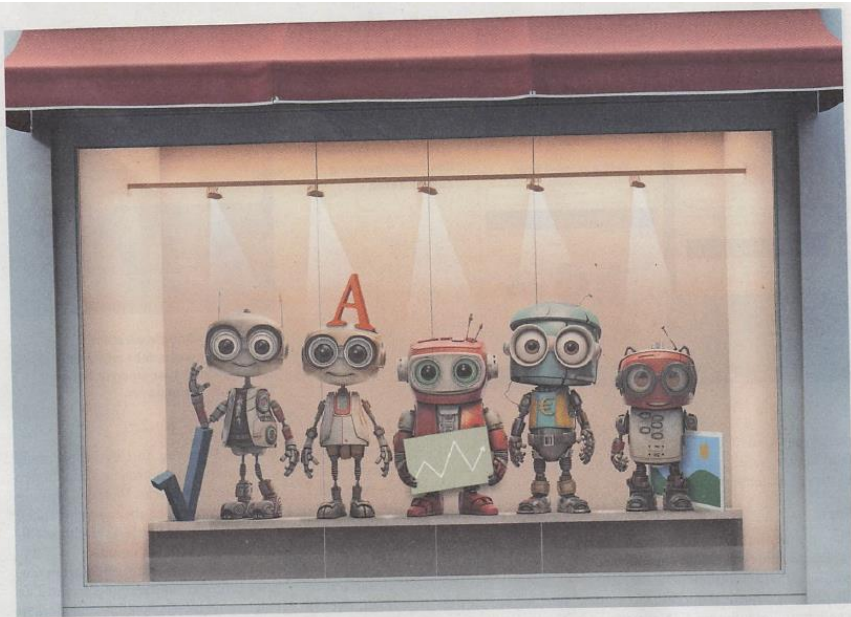
Zusammenfassung / Fazit:

- Die Generierung des Ergebnisses ist nicht reproduzierbar! ChatGPT wählt häufig unterschiedliche Strategien, um die gestellte Aufgabe zu erledigen... oder weigert sich ganz, den Prompt auszuführen.
- Nach unseren Erfahrungen scheint das Training/finetuning des GPT die Ergebnisse zu verbessern, insbesondere was die korrekte Annotation der Personen-Orte Relation betrifft.
- Es fällt ChatGPT schwer, ein Itinerar auf einer Karte darzustellen. Es werden sehr unterschiedliche Visualisierungen gewählt und die Bitte um Nummerierung der Orte nach dem Itinerar wird häufiger auch ganz ignoriert.



Zusammenfassung / Fazit:

- Insgesamt betrachtet haben LLMs gute Fähigkeiten im Bereich von NER und eignen sich damit für die Erkennung und Auszeichnung von Orten und Personen.
- Größere Textmengen können bislang nur über einen (kostenpflichtigen) API Zugang zum LLM verarbeitet werden. Bislang ist der Umfang der automatisch annotierbaren Textmenge im Chat sehr begrenzt (einige Seiten).
- Ein trainiertes Custom GPT kann eine manuelle Annotation unterstützen und evtl. sogar ersetzen, aber nicht ohne eine genaue Kontrolle (und ggf. Korrektur) des Ergebnisses!



Bot-Marktplatz

Maßgeschneiderte ChatGPT-Bots im Feldversuch

ChatGPT-Nutzer können seit Kurzem individuelle Chatbots, sogenannte GPTs, bauen und anderen zur Verfügung stellen.

Fragt man ChatGPT danach, wofür er sich nicht gut eignet, antwortet der Chatbot unter anderem: „In sehr spezialisierten oder aktuellen wissenschaftlichen Bereichen...“

ten – würden zum Beispiel einen prima Support-Bot ergeben.

Die im November vorgestellten GPTs können solche Aufgaben übernehmen. GPT steht wie bei ChatGPT für „Generative Pre-trained Transformer“, eine bestimmte Architektur künstlicher neuronaler Netze – und für sich. GPTs sind von Benutzern erstellte Chatbots auf Basis von ChatGPT, die diese auf bestimmte Aufgaben zugeschnitten haben.

Zusätzlich zum Allgemeinwissen von ChatGPT können sie daher auch umfassende Kenntnisse zu speziellen Themen haben.

Fazit

Individuelle Versionen von ChatGPT bauen, mit einer eigenen Wissensbasis und maßgeschneiderten Funktionen: hört sich gut an. Probiert man GPTs aus, stellt sich aber schnell Ernüchterung ein: ein unübersichtlicher Store, viele schrottige Bots, ungeschützte Prompts, mangelhafte Dokumentation, beschränkte Funktionen, Aussetzer und vieles mehr. Derzeit hinterlassen die GPTs eher den Eindruck eines frühen Machbarkeitstests. Experimentieren kann man damit ein wenig, produktiv sollte man sie nicht einsetzen.

(jo@ct.de) **ct**

GPT-Verzeichnisse: ct.de/y4f8



Vielen Dank für Ihre Aufmerksamkeit!

Custom-GPTs:

<https://chat.openai.com/g/g-M46uUIBD6-travelogueannotator>

<https://chat.openai.com/g/g-Uy5ZolqD2-itinerizer>



**Göttinger
Digitale
Akademie**

Tarjia Alam Nisha
Georg-August-Universität Göttingen
tarjia.alamnisha@stud.uni-goettingen.de

Dr. Franziska Pannach
Rijksuniversiteit Groningen
f.a.pannach@rug.nl

Dr. Jörg Wettlaufer
Niedersächsische Akademie der Wissenschaften zu Göttingen
joerg.wettlaufer@adwgoe.de