

- openHPI: ChatGPT: Was bedeutet generative KI für unsere Gesellschaft? -

Prototyping und Datentransformation

Johannes Hötter
Christian Warmuth

Szenario: Chatbot für Buchung von Flügen

Ziel

- Proprietäres Chat-Modell bauen, welches Useranfragen rund um Flugbuchungen beantworten kann
- Modell soll “in-house” laufen können

Problem

- Noch keine Daten vorhanden für eigenes Modell

Idee

- ChatGPT hat “Weltwissen”, was auch Flüge beinhaltet



Prototyping und Transformationen

Johannes Hötter,
Christian Warmuth

openHPI


Szenario: Chatbot für Buchung von Flügen


↺ Clear chat

Chat History is off.
Chats won't be saved in your history or used for training our models to improve ChatGPT. Unsaved chats will be deleted from our systems within 30 days. [Learn more](#)

⏻ Enable chat history

Model: GPT-4


 Stelle dir vor, Du bist ein Kunde einer Fluggesellschaft. Schreibe 5 Fragen, die etwa an den Service Desk gehen könnten. Erzeuge ein gültiges JSON-File dafür.



```
{  
  "fragen": [  
    {  
      "frage_id": 1,  
      "frage": "Kann ich meinen Flug umbuchen oder stornieren, und welche Gebühren fallen dabei an?"  
    },  
    {  
      "frage_id": 2,  
      "frage": "Wie viel Gepäck darf ich kostenfrei mitnehmen, und welche Gebühren fallen für Übergepäck an?"  
    },  
    {  
      "frage_id": 3,  
      "frage": "Wie kann ich meinen Sitzplatz im Voraus reservieren, und welche Kosten sind damit verbunden?"  
    },  
    {  
      "frage_id": 4,
```

Regenerate response

Send a message.

 Johannes Hötter

...

ChatGPT may produce inaccurate information about people, places, or facts. [ChatGPT May 3 Version](#)

Prototyping und Transformationen

Johannes Hötter,
Christian Warmuth

openHPI

Szenario: Chatbot für Buchung von Flügen

Rahmenbedingungen

- ChatGPT kann nur Trainingsdaten zu Themen generieren, die es kennt
- Immer im Blick behalten, ob die AGBs von OpenAI dafür eingehalten werden; siehe AGBs (siehe nächste Folie)
- Natürlich nach wie vor Datenqualität im Auge behalten
- **Keine sensiblen Daten eingeben!**

Vorteile

- Sehr kostengünstig
- Kann bestehende Datenbasis prima ergänzen



Prototyping und Transformationen

Johannes Hötter,
Christian Warmuth

openHPI

AGBs von OpenAI: Usage Requirements



[Research](#) ▾ [Product](#) ▾ [Developers](#) ▾ [Safety](#) [Company](#) ▾

2. Usage Requirements

(a) **Use of Services.** You may access, and we grant you a non-exclusive right to use, the Services in accordance with these Terms. You will comply with these Terms and all applicable laws when using the Services. We and our affiliates own all rights, title, and interest in and to the Services.

(b) **Feedback.** We appreciate feedback, comments, ideas, proposals and suggestions for improvements. If you provide any of these things, we may use it without restriction or compensation to you.

(c) **Restrictions.** You may not (i) use the Services in a way that infringes, misappropriates or violates any person's rights; (ii) reverse assemble, reverse compile, decompile, translate or otherwise attempt to discover the source code or underlying components of models, algorithms, and systems of the Services (except to the extent such restrictions are contrary to applicable law); (iii) use output from the Services to develop models that compete with OpenAI; (iv) except as permitted through the API, use any automated or programmatic method to extract data or output from the Services, including scraping, web harvesting, or web data extraction; (v) represent that output from the Services was human-generated when it is not or otherwise violate our Usage Policies; (vi) buy, sell, or transfer API keys without our prior consent; or (viii), send us any personal information of children under 13 or the applicable age of digital consent. You will comply with any rate limits and other requirements in our documentation. You may use Services only in geographies currently supported by OpenAI.

(d) **Third Party Services.** Any third party software, services, or other products you use in connection with the Services are subject to their own terms, and we are not responsible for third party products.

Quelle: <https://openai.com/policies/terms-of-use>

Stand: 07.05.2023

GEFÖRDERT VOM



Bundesministerium
für Bildung
und Forschung

Prototyping und Transformationen

Johannes Hötter,
Christian Warmuth

openHPI

Korrektur von Texten mit Fehlern, z.B: OCR



Correct any mistakes in this paragraph:



This agreeinent ("Agreeneent") is entered into on this 7th day of Way 2023 by and between [Compnay Name], a corparation orgenized and exesting under the laws of the Srate of Califomia, with its principal place of busniness at [Adress], and [Cilent Name], an individuel residing at [Address]. The parties hereby agree to engage in a bussiness relationship for the purpose of [Purpose of Agreeemnt]. This Agreement shall comence on the date of excecution by both parties and shall continne until terminated by either party in accordance with the provisions of this Agreement. The terms and conditoins of this Agreement are as follows: [Instert Terms and Condditions].



This agreement ("Agreement") is entered into on this 7th day of May 2023 by and between [Company Name], a corporation organized and existing under the laws of the State of California, with its principal place of business at [Address], and [Client Name], an individual residing at [Address]. The parties hereby agree to engage in a business relationship for the purpose of [Purpose of Agreement]. This Agreement shall commence on the date of execution by both parties and shall continue until terminated by either party in accordance with the provisions of this Agreement. The terms and conditions of this Agreement are as follows: [Insert Terms and Conditions].



Prototyping und Transformationen

Johannes Hötter,
Christian Warmuth

openHPI

Identifikation von Dopplungen



I have two customers:

- 12.04.2020: John Doe, 29 years old, living at Fürstenstraße 5, Bonn, with 3 kids
- 02.09.2023: John Doe, 42 years old, living in Stahnsdorfer Straße 144, Potsdam, with 2 kids

Is this likely the same person? Reason why.



It is unlikely that these two customers are the same person, despite sharing the same name. Here are the reasons why:

1. Age discrepancy: The first John Doe is 29 years old on April 12, 2020, while the second John Doe is 42 years old on September 2, 2023. If they were the same person, the first John Doe would be 32 years old in 2023, not 42 years old. This is a significant age difference.
2. Address difference: The first John Doe lives in Bonn, while the second one lives in Potsdam. While it's possible for someone to move between these cities, the age discrepancy makes it less likely that this is the case.
3. Number of kids: The first John Doe has 3 kids, while the second one has 2 kids. It is possible that the number of kids is incorrect in one of the records, but combined with the other discrepancies, this further supports the idea that these are two different people.

Prototyping und Transformationen

Johannes Hötter,
Christian Warmuth

openHPI

Identifikation von Dopplungen



I have two customers:

- 12.04.2020: John Doe, 39 years old, living at Fürstenstraße 5, Bonn, with 2 kids
- 02.09.2023: John Doe, 42 years old, living in Stahnsdorfer Straße 144, Potsdam, with 3 kids

Is this likely the same person? Reason why.



While it is not possible to definitively confirm whether the two John Does are the same person, there are certain factors that suggest they might be. Here are the key similarities and differences between the two:

Similarities:

1. Same name: Both individuals are named John Doe.
2. Age progression: The age difference between the two John Does is three years, which aligns with the difference in the dates provided (April 2020 to September 2023).

Differences:

1. Addresses: The two individuals live in different cities, with the first John Doe residing in Bonn, and the second in Potsdam.
2. Number of kids: The first John Doe has 2 kids, while the second has 3 kids.

Based on these factors, it is plausible that the two customers could be the same person. The age progression is consistent, and it's possible that this individual moved from Bonn to Potsdam during the intervening years and had an additional child. However, without further information or verification, it's impossible to say for certain.

**Prototyping und
Transformationen**

Johannes Hötter,
Christian Warmuth

openHPI

Identifikation von Dopplungen



Give a likelihood score from 0 to 100 for them being the same person



Given the information provided, I would assign a likelihood score of 70 out of 100 that they are the same person. This score reflects the reasonable possibility that these two customers are the same individual, considering the consistent age progression and the potential for life changes like moving and having another child. However, the score isn't higher because there is still some uncertainty due to the differences in address and number of kids, and the commonality of the name John Doe.

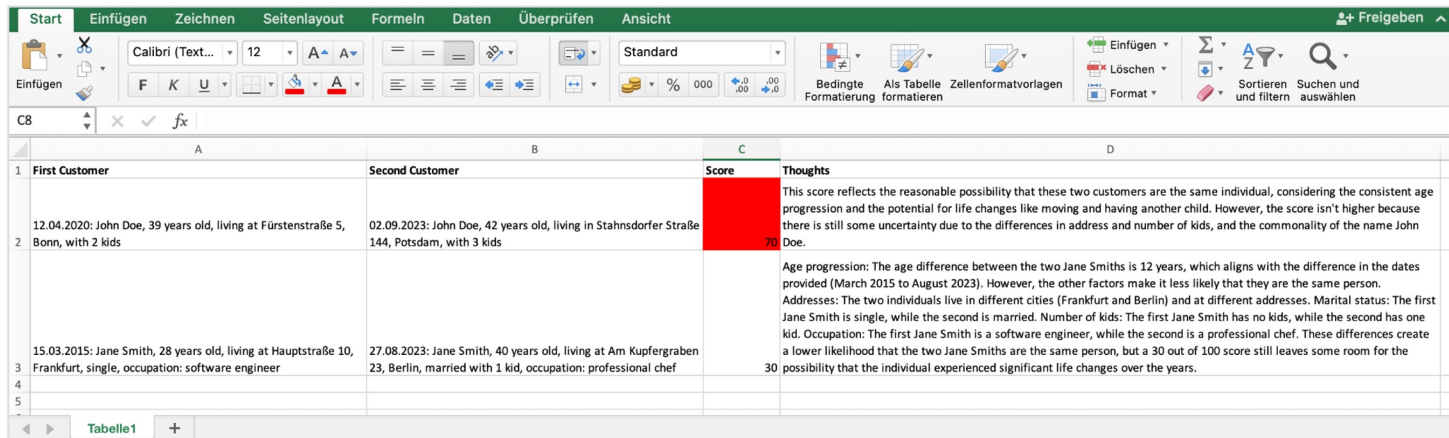


Prototyping und Transformationen

Johannes Hötter,
Christian Warmuth

openHPI

Human-in-the-loop



	A	B	C	D
1	First Customer	Second Customer	Score	Thoughts
2	12.04.2020: John Doe, 39 years old, living at Fürstenstraße 5, Bonn, with 2 kids	02.09.2023: John Doe, 42 years old, living in Stahnsdorfer Straße 144, Potsdam, with 3 kids	70	This score reflects the reasonable possibility that these two customers are the same individual, considering the consistent age progression and the potential for life changes like moving and having another child. However, the score isn't higher because there is still some uncertainty due to the differences in address and number of kids, and the commonality of the name John Doe.
3	15.03.2015: Jane Smith, 28 years old, living at Hauptstraße 10, Frankfurt, single, occupation: software engineer	27.08.2023: Jane Smith, 40 years old, living at Am Kupfergraben 23, Berlin, married with 1 kid, occupation: professional chef	30	Age progression: The age difference between the two Jane Smiths is 12 years, which aligns with the difference in the dates provided (March 2015 to August 2023). However, the other factors make it less likely that they are the same person. Addresses: The two individuals live in different cities (Frankfurt and Berlin) and at different addresses. Marital status: The first Jane Smith is single, while the second is married. Number of kids: The first Jane Smith has no kids, while the second has one kid. Occupation: The first Jane Smith is a software engineer, while the second is a professional chef. These differences create a lower likelihood that the two Jane Smiths are the same person, but a 30 out of 100 score still leaves some room for the possibility that the individual experienced significant life changes over the years.
4				
5				

Vorgehen:

- Integriertes LLM (z.B. open-source) prognostiziert paarweise Dopplungen, dann z.B. per Excelmappe exportieren
- Prüfen von Einträgen mit hohem Score

Prototyping und Transformationen

Johannes Hötter,
Christian Warmuth

openHPI

- openHPI: ChatGPT: Was bedeutet generative KI für unsere Gesellschaft? -

Prototyping und Datentransformation

Johannes Hötter
Christian Warmuth