



GENAI FÜR DIE PRAXIS

CorrelCon 2024 | Michael Aydinbas



MICHAEL AYDINBAS

Data Engineer

DRIVE POSITIVE IMPACT

- MSc Psychology in IT, TU Darmstadt
- 4 years as consultant at Exxeta
- Senior Data Engineer at New Work SE
- Project Manager at CorrelAid since 2020
- RLernen & Data Literacy tutor
- Trustee & volunteer Chief Data Officer at Action4Diabetes
- Lecturer, Author

TODAY'S WORKSHOP

In the first part of the session, I will introduce the project.

In the larger second part of the session, I will demonstrate our approach.

At the end I hope for a lively discussion about the benefits and drawbacks of using genAI.

The Project

01

Getting started with LLMs

02

Workshop

03

Discussion

04

A detailed landscape painting featuring a river flowing through a lush green valley. In the foreground, a large, gnarled tree stands on the left, and another smaller tree is on the right. The river is filled with rocks and has white water rapids. In the background, majestic mountains with sharp peaks rise against a sky with soft, white clouds. The overall mood is serene and natural.

THE PROJECT

BABYLOTSE FRANKFURT

- The program is offered by the Deutscher Kinderschutzbund, Bezirksverband Frankfurt e.V. and reaches more than 13,000 families per year, of which about 35% take advantage of the personal consultation.
- Babylotse has conducted a small survey among parents in Frankfurt. The data collected is very unstructured due to the open-ended nature of the questions, and there are also some spam responses in the data set.
- Babylotse is looking for support to analyze the data. The goal is to identify the most important "tops" and "flops" (see survey) and compile them in a list, possibly broken down by neighborhood.
- In this way, Babylotse will be able to send a clear and concise message to political representatives about what needs to be done in Frankfurt to make the city more family friendly.
- The challenge is in cleaning the data
- Since we are dealing with anonymized data, the idea is to implement data cleansing and analysis using ChatGPT or a similar LLM.
- Steps:
 - Identify categories (tops & flops)
 - For example, up to 5 categories per sentiment
 - Assign the answers to the categories
 - Per category e.g. the 3 TOP reasons
- Filter out spam responses
- Simple descriptive analysis (tables, possibly bar charts) in consultation with Babylotse

APPROACH

This is a very high-level picture, more details later in the workshop.

There are more preprocessing steps as well as a mapping of free-text answers to the corresponding city neighborhood.



**Find
categories**



**Improve
categories**



**Classify
comments**



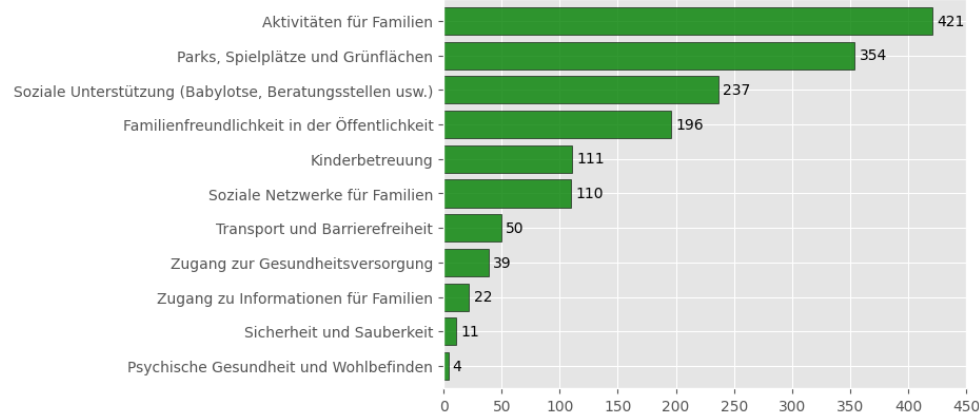
**Analyze
comments**

Refine

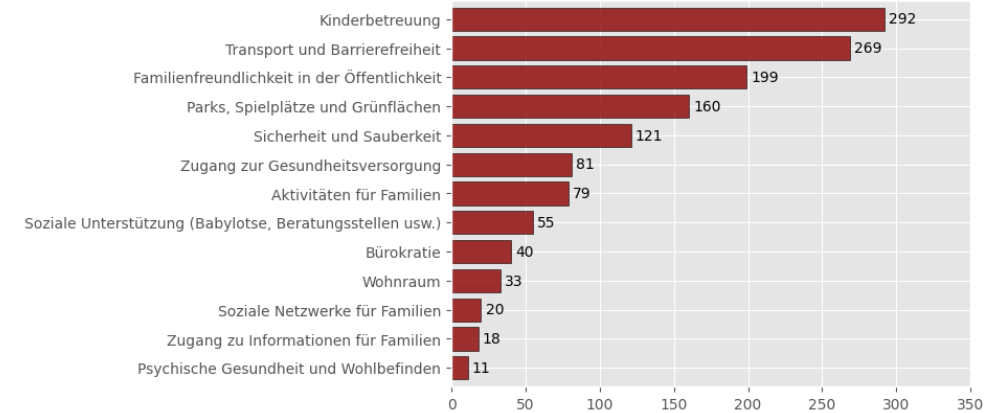
GLOBAL ANALYSIS

RESULTS

Positive Aspekte über das Familienleben mit Baby in Frankfurt



Negative Aspekte über das Familienleben mit Baby in Frankfurt



Frequency of each positive category

One comment can be assigned to multiple categories

Frequency of each negative category

One comment can be assigned to multiple categories

REGIONAL ANALYSIS

RESULTS

			Positive
Region	category		
Mitte	Aktivitäten für Familien		83
	Parks, Spielplätze und Grünflächen		81
	Soziale Unterstützung (Babylotse, Beratungsstellen usw.)		62
Nord	Aktivitäten für Familien		126
	Parks, Spielplätze und Grünflächen		111
	Familienfreundlichkeit in der Öffentlichkeit		70
Ost	Aktivitäten für Familien		94
	Parks, Spielplätze und Grünflächen		72
	Soziale Unterstützung (Babylotse, Beratungsstellen usw.)		50
Süd	Aktivitäten für Familien		48
	Parks, Spielplätze und Grünflächen		44
	Soziale Unterstützung (Babylotse, Beratungsstellen usw.)		35
Umland	Aktivitäten für Familien		5
	Familienfreundlichkeit in der Öffentlichkeit		5
	Soziale Unterstützung (Babylotse, Beratungsstellen usw.)		5
West	Aktivitäten für Familien		64
	Parks, Spielplätze und Grünflächen		42
	Familienfreundlichkeit in der Öffentlichkeit		24

Top 3 positive categories per region

			Negative
Region	category		
Mitte	Transport und Barrierefreiheit		68
	Kinderbetreuung		58
	Familienfreundlichkeit in der Öffentlichkeit		53
Nord	Kinderbetreuung		83
	Transport und Barrierefreiheit		74
	Familienfreundlichkeit in der Öffentlichkeit		49
Ost	Kinderbetreuung		61
	Transport und Barrierefreiheit		52
	Familienfreundlichkeit in der Öffentlichkeit		48
Süd	Kinderbetreuung		48
	Transport und Barrierefreiheit		39
	Familienfreundlichkeit in der Öffentlichkeit		19
Umland	Transport und Barrierefreiheit		5
	Bürokratie		3
	Familienfreundlichkeit in der Öffentlichkeit		2
West	Kinderbetreuung		39
	Parks, Spielplätze und Grünflächen		34
	Transport und Barrierefreiheit		31

Top 3 negative categories per region

CATEGORY ANALYSIS

RESULTS

AI generated

Category: Childcare and Early Education Resources

Topic 1: **Insufficient Availability of Childcare Places**

This topic revolves around the overwhelming concern regarding the lack of available spots in kindergartens and childcare facilities (Kitas and Krippen). Many comments express frustration about the difficulty in securing a place for children aged 0-5, with numerous respondents highlighting that they have been searching for a spot for an extended period without success. The comments indicate a widespread sentiment that the current number of available places is inadequate to meet the demand.

Topic 2: **Staffing Shortages and Quality of Care**

Another significant issue raised in the comments is the shortage of qualified staff in childcare facilities. Respondents mention that this staffing crisis leads to reduced quality of care, frequent closures, and limited operational hours. Many express concern that the lack of personnel negatively impacts the overall experience and care provided to children, contributing to a sense of instability in the childcare environment.

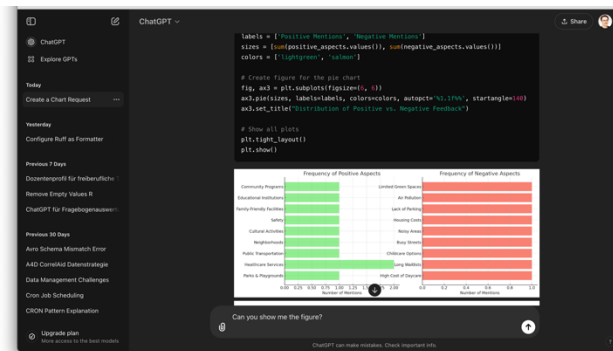
...

A detailed landscape painting featuring a river flowing through a lush green valley. In the foreground, a large, gnarled tree stands on the left, and another smaller tree is on the right. The river is filled with rocks and has white water rapids. In the background, majestic mountains with sharp peaks rise against a sky with soft, white clouds. The overall tone is serene and naturalistic.

GETTING STARTED WITH LLM

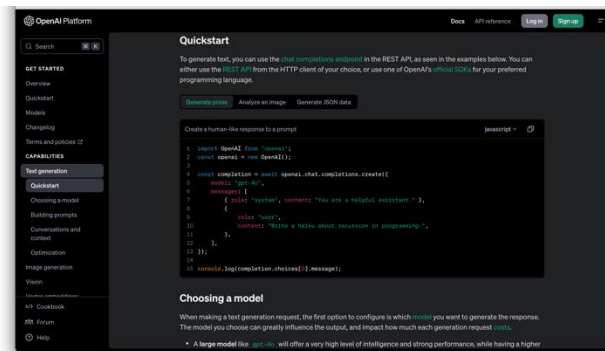


ChatGPT



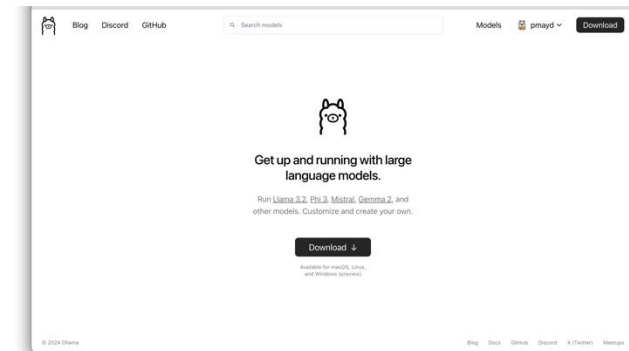
<https://chatgpt.com/share/6706f178-6d78-8008-903d-5ae5ec114db0>

OpenAI API



<https://github.com/openai/openai-python?tab=readme-ov-file#usage>
OSS:
<https://console.groq.com/docs/quickstart>

Local & OSS (e.g. Ollama)



<https://ollama.com/>
<https://github.com/open-webui/open-webui>

GET YOUR DEVELOPER API KEY

- Need to register
- Create project: <https://platform.openai.com/organization/projects>
- Create keys: <https://platform.openai.com/api-keys>
- Cost control: <https://platform.openai.com/usage>
- Always costs: <https://openai.com/api/pricing/>

OpenAI

- Need to register
- Create key: <https://console.groq.com/keys>
- Cost control: <https://console.groq.com/settings/usage>
- Actually, free as long as you stay within the limits and don't upgrade
- Limits apply: <https://console.groq.com/settings/limits>

Grog Cloud

A detailed landscape painting featuring a river flowing through a lush green valley. In the foreground, a large, gnarled tree stands on the left, and another with autumn-colored leaves is on the right. The river is filled with rocks and rapids. In the background, majestic mountains with snow-capped peaks rise against a cloudy sky. A semi-transparent teal rectangle is centered over the image, containing the word 'WORKSHOP' in white capital letters.

WORKSHOP

GITHUB REPOSITORY

<https://github.com/CorrelAid/workshop-babylotse>

A detailed landscape painting featuring a river flowing through a lush green valley. In the foreground, a large, gnarled tree stands on the left, and another smaller tree is on the right. The river is filled with rocks and has a small waterfall. In the background, majestic mountains with sharp peaks rise against a sky with soft, golden clouds. The overall mood is serene and natural.

DISCUSSION

LEARNINGS & INSIGHTS

LLMs are versatile

Low or no data pre-processing

Very large context windows



Struggles with short text

Batch API only for the patient

Human-in-the-loop approach