

>> neue fische
School and Pool for Digital Talent

Code of Conduct

We expect all participants to our events and community to abide to this code of conduct:
LadyNerds Code of Conduct (<http://bit.ly/LadyNerds-CoC>).

We follow the **LadyNerds Code of Conduct** because we are dedicated to providing a safe, inclusive, welcoming, and harassment-free space and experience for all members and guests, regardless of gender identity and expression, sexual orientation, disability, physical appearance, socioeconomic status, body size, ethnicity, nationality, level of experience, age, or religion (or lack thereof).

The Code of Conduct exists because of that dedication. We do not tolerate harassment in any form and we prioritise marginalised people's safety over privileged people's comfort.



Map

neue fische

- 8 Bootcamp Programs
- 5 Locations
- 900+ Graduates
- 85% Job success



Quelle: ¹Bitkom, ²Future of Jobs Report 2020, ³Studie EY(2016)

neuefische.de



Data Bootcamps coming up:

- 20 June: Data Practitioner Part Time - 24 weeks
- 26 June: Machine Learning Engineer - 4 weeks
- 17 July: Data Analytics - 12 weeks
- 24 July: Data Science - 12 weeks

22 June: workshop: A Practical approach to Clustering

How to get in touch:

email: studienberatung@neuefische.de

Unleash the Power of LLMs: Crack the Code of Perfect Prompts

with Kavish Bhatia and Andreas Dosch



About me

Kavish Bhatia
Data Science Coach

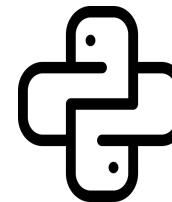
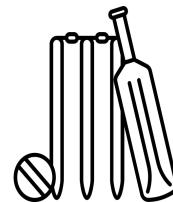


My background:

Bachelors in Computer Science
Masters in Machine Learning

Ask me about:

Football, Cricket, Automating stuff



kavish.bhatia@neuefische.de

About me

Andreas Dosch

Acting Boot Camp manager in the
Data Practitioner course



My background:

Physicist & science school teacher

Ask me about:

My family, meditation, guitar



andreas.dosch@neuefische.de

Agenda

- Introduction
- Prompt Engineering Basics
- GPT-4 & Plugin Workflows
- Prompting with OpenAI and LangChain
- Evaluating Model Response
- Summary
- Q&A

Philosophenweg (Philosophers' Walk)

Soon in English!

Wünschelrute

Schläft ein Lied in allen Dingen,
die da träumen fort und fort
und die Welt hebt an zu singen,
triffst du nur das **Zauberwort.**



Joseph Freiherr von Eichendorff (Wikipedia, H.-P.Haack)

Wishing-Wand

A song sleeps in all things around
Which dream on and on unheard,
And the world begins to resound,
If you hit the **magic word.**

Translated by Natas Neutert



Joseph Freiherr von Eichendorff (Wikipedia, H.-P.Haack)

Brute Force
Do **ten more**, pick to your taste!

All things possess a drowsing rhyme,
Dreaming away in the hands of time,
And the world's song begins to lift, once
the magic word unveils its gift.

In all things a song slumbers deep,
Continuously dreaming as they weep,
The world sings when you discover,
The magic word that serves as a cover.

A tune rests in every thing,
Dreams unfold on constant wing,
The world shall hum its song anew,
Once the magic word is found by you.

Within all things a lullaby hides,
Dreaming incessantly as time bides, The
world starts singing in unison, When you
unveil the magic that's spun.

In every object, a song lies asleep,
Dreaming endlessly, a secret to
keep, The world starts to sing a
wondrous chord, When the magic
word you finally record.

All things carry a sleeping tune,
Dreaming onwards, beneath the
moon, The world sings with a heart
beguiled, When the magic word is
reconciled.

A slumbering song in all things
dwell, Endless dreaming, a secret
spell, The world bursts into song so
bright, When the magic word is
brought to light.

Each object holds a dormant lay,
Dreaming on through night and
day, When the magic word you
chance to learn, The world's
harmonious song shall return.



Do ten
more!

Inside everything, a song does rest,
In an eternal dream, the notes are
best, The world sings when you
scome to find, The magic word that's
been intertwined.

Joseph Freiherr von Eichendorff.

A hushed melody within all things
Joseph Freiherr von Eichendorff (Wikipedia)
resides, In dreams that forever, like
rivers, glides, And the world
awakens in song so pure, When the
magic word you at last secure.

Agenda

- Introduction
- **Prompt Engineering Basics**
- GPT-4 & Plugin Workflows
- Prompting with OpenAI and LangChain
- Evaluating Model Response
- Summary
- Q&A

Leverage the mere fact
that GPT is *not* you and get it's opinion

1. Act as a (harsh) critic and provide (brutally honest) feedback about these {{email titles}}.
Convince me why they are bad.
2. Out of all {{those}}, which would you pick?
Rewrite five variations and convince me, why these are better.

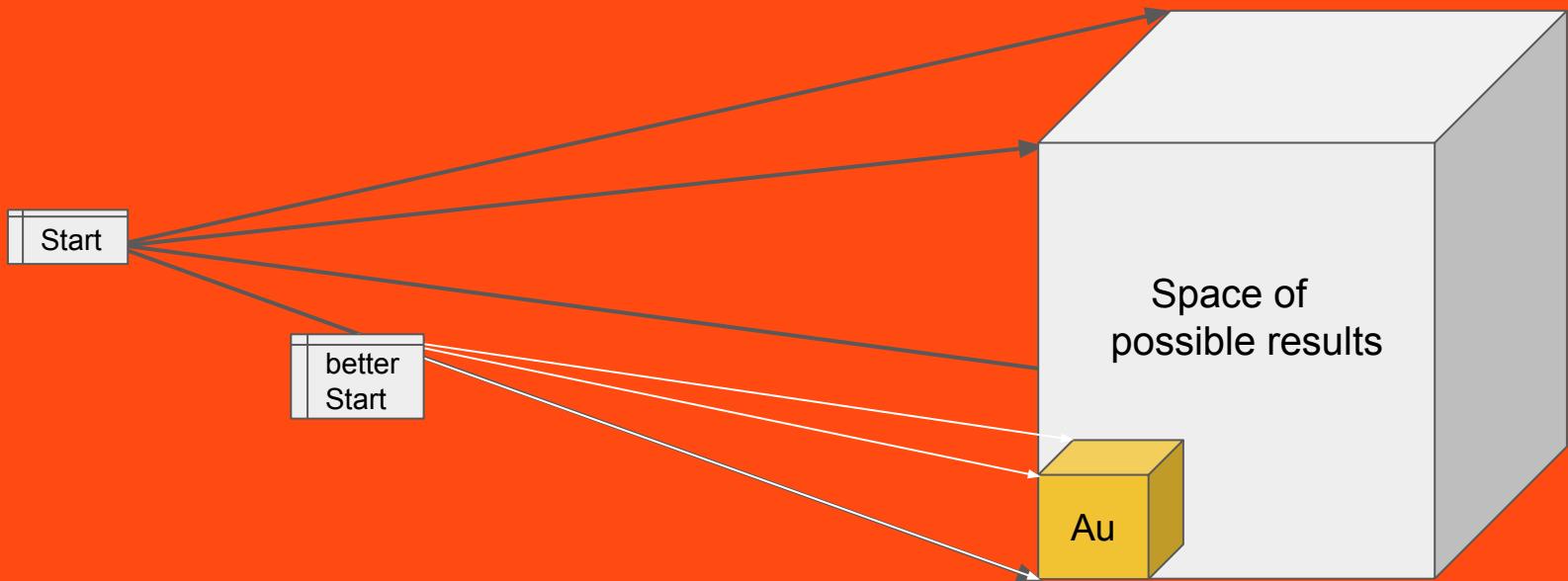
A little Prompt Engineering Philosophy

This is
where the slides
for the
all-over-the-internet-stuff
were

→ show notes

Start

Space of
possible results



Conclusion: 4 Tips for Developing A Prompt Engineering Mindset

- 1. Describe like it already exists**
- 2. Provide Instructions and Guidance, not a Formula**
- 3. Think and Describe in Analogies**
- 4. “Genius” in a room**

Agenda

- Introduction
- Prompt Engineering Basics
- **GPT-4 & Plugin Workflows**
- Prompting with OpenAI and LangChain
- Evaluating Model Response
- Summary
- Q&A

remember that I play the guitar?

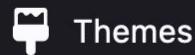
**→ the amp is considered
part of the instrument!**

Let's talk about *plugins*

browser-plugins ≠ GPT-plugins



Extensions



Themes



Available
Updates

1



Superpower ChatGPT



...

ChatGPT with superpowers! Sync/search history locally, create folders, export all chats, pin messages, access thousands of prompts

Details

Permissions

Release Notes



Wall of text

Show more

Settings

Close

Your can change the Superpower settings here

General

Auto Sync

models

Custom Prompts

Export

Splitter

Newsletter

Custom Prompts

Add New Custom Prompts

@Structure

`role`:__ `result`:__ `context/content`:__ `goal`:__

Delete

Default

@Markdown

Display the results in Markdown. GPT can do lists! If you want lists, try my buddy @Lists

Delete

Default

@Custom_2

It's me, your OTHER favourite prompt puzzle piece

Delete

Default

@Custom_1

It's me, your favourite prompt puzzle piece 1

Delete

Default

@Continue

Please continue

Delete

Default



Powered by [Superpower ChatGPT](#) (v 3.6.0 [Release Note](#))

Made by [Saeed Ezzati](#) - [Support this extension ➔](#)

+ New chat

Model

Plugins Beta gpt4

Search conversations



Fine-tuning LLM Offline

06/10/23 11:28 PM

prompt perfect showcase

06/09/23 5:44 PM

All that prompting

NYC Taxi Data 2022

05/25/23 10:27 PM

CLI terminal commands

05/25/23 9:15 PM

SQL

New chat

05/28/23 10:30 PM

upload auto

Experimente berufl

Experimente privat

Docker for Data Science

My Prompt History

Community Prompts

Plugins

Beta gpt4

Plugins



3/3 Enabled

PlugFinder

Prompt Perfect

AskYourPDF

Access Link

Noteable

Show Me Diagrams

edX

Link Reader

ScholarAI

Ambition

Video Insights

Tone

Confident

Writing Style

Descriptive

Language

English



be used to improve

be deleted from our

⋮ Continue ⏪

Send a message.



0 chars / 0 words

GPT4 requests (last 3 hours): 0/25



+ New chat

Model

Plugins Beta gpt4

Search conversations



Fine-tuning LLM Offline

06/10/23 11:28 PM

prompt perfect showcase

06/09/23 5:44 PM

All that prompting

NYC Taxi Data 2022

05/25/23 10:27 PM

CLI terminal commands

05/25/23 9:15 PM

SQL

New chat

05/28/23 10:30 PM

upload auto

Experimente berufl

Experimente privat

Docker for Data Science

My Prompt History

Community Prompts



Synced

Plugins



Tone

Confident

Writing Style

Descriptive

Language

English

3/3 Enabled

PlugFinder

Prompt Perfect

AskYourPDF

Access Link

Noteable

Show Me Diagrams

edX

Link Reader

ScholarAI

Ambition

Video Insights



LUS

be used to improve
be deleted from our

⋮ Continue ⏪

Send a message.



GPT4 requests (last 3 hours): 0/25

0 chars / 0 words



+ New chat

Model

Plugins Beta gpt4

Search conversations



Fine-tuning LLM Offline

06/10/23 11:28 PM

prompt perfect showcase

06/09/23 5:44 PM

All that prompting

NYC Taxi Data 2022

05/25/23 10:27 PM

CLI terminal commands

05/25/23 9:15 PM

SQL

New chat

05/28/23 10:30 PM

upload auto

Experimente berufl

Experimente privat

Docker for Data Science

My Prompt History

Community Prompts



Synced

Plugins

Beta gpt4

Plugins



3/3 Enabled

PlugFinder

Prompt Perfect

AskYourPDF

Access Link

Noteable

Show Me Diagrams

edX

Link Reader

ScholarAI

Ambition

Video Insights

Tone

Confident

Writing Style

Descriptive

Language

English



be used to improve

be deleted from our

⋮ Continue ⏪

Send a message.



GPT4 requests (last 3 hours): 0/25

0 chars / 0 words



+ New chat

Model
Plugins Beta gpt4

Search conversations

Fine-tuning LLM Offline
06/10/23 11:28 PM

prompt perfect showcase
06/09/23 5:44 PM

All that prompting

NYC Taxi Data 2022
05/25/23 10:27 PM

CLI terminal commands
05/25/23 9:15 PM

SQL

New chat
05/28/23 10:30 PM

upload auto

Experimente berufl

Experimente privat

Docker for Data Science

My Prompt History

Community Prompts

Synced

Plugins



Tone

Confident

Writing Style

Descriptive

Language

English

3/3 Enabled

PlugFinder	<input type="radio"/>
Prompt Perfect	<input checked="" type="radio"/>
AskYourPDF	<input checked="" type="radio"/>
Access Link	<input type="radio"/>
Noteable	<input checked="" type="radio"/> LUS
Show Me Diagrams ❌	<input type="radio"/>
edX	<input type="radio"/>
Link Reader ❌	<input type="radio"/>
ScholarAI	<input type="radio"/>
Ambition	<input type="radio"/>
Video Insights	<input type="radio"/>

be used to improve
be deleted from our

Send a message.



GPT4 requests (last 3 hours): 0/25

0 chars / 0 words

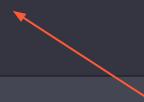


SH

Write the source code for the next "facebook"



Perfect



< 2/2 >



Used **Prompt Perfect**



REQUEST TO PROMPT PERFECT



```
{  
  "text": "Write the source code for the next 'facebook'"  
}
```

RESPONSE FROM PROMPT PERFECT



```
{  
  "rephrased": {  
    "text": "Please write the complete and detailed source code for a social  
networking website that is comparable to Facebook in terms of functionality and  
features. The code should be designed to handle a large number of users and be  
scalable for future growth. It should also include appropriate security measures to  
protect user data and privacy. Please provide clear documentation and comments to aid  
in understanding and maintenance of the code."  
  }  
}
```

→ let's start a
new conversation
with GPT



Model: Plugins • Enabled plugins:

SH

You are a Data Science coach.

Use the content of <https://docker-curriculum.com/> to compile a notebook for data science bootcamp participants with some experience.





SH

You are a Data Science coach.

Use the content of <https://docker-curriculum.com/> to compile a notebook for data science bootcamp participants with some experience.



Used **Access Link**



The content from the provided URL is about Docker, which is a platform that allows users to automate the deployment, scaling, and management of applications. Docker is not specifically a data science tool, but it is often used in data science environments to ensure consistent and reproducible results across different machines.

Here's a brief outline of how we can structure the notebook:

1. **Introduction to Docker**

- What is Docker?
- Why is Docker important for data science?

2. **Installing Docker**



Model: Plugins • Enabled plugins:

SH

You are a Data Science coach.

Use the content of <https://docker-curriculum.com/> to compile a notebook for data science bootcamp participants with some experience.



Used **Access Link**



REQUEST TO ACCESS LINK



```
{  
  "url": "https://docker-curriculum.com/",  
  "chat_context": "The user has asked me to compile a notebook for data science  
bootcamp participants with some experience using the content of the provided URL."  
}
```

RESPONSE FROM ACCESS LINK



```
"Prakhar Srivastav\n\nINTRODUCTION\n\nWHAT IS DOCKER?\nWikipedia defines Docker  
as\nan open-source project that automates the deployment of software applications  
inside containers by providing an additional layer of abstraction and automation of  
OS-level virtualization on Linux.\nWow! That's a mouthful. In simpler words, Docker  
is a tool that allows developers, sys-admins etc. to easily deploy their applications  
in a sandbox (called containers) to run on the host operating system i.e. Linux.
```

Here's a brief outline of how we can structure the notebook:

- 1. Introduction to Docker**

- What is Docker?
- Why is Docker important for data science?

- 2. Installing Docker**

- Instructions for different operating systems

- 3. Docker Basics**

- Docker Images
- Docker Containers
- Docker Commands

- 4. Building a Docker Image**

- Dockerfile
- Building an Image

- 5. Docker Compose**

- What is Docker Compose?
- Docker Compose YAML file
- Running multi-container Docker applications

- 6. Docker in Data Science**

- Using Docker for reproducibility

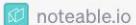
Let's start creating the notebook using the Noteable plugin.

Used Noteable



Great! I've created a new notebook named "Docker_for_Data_Science.ipynb". You can follow along with the notebook [here](#).

Now, let's start adding content to the notebook. We'll begin with the introduction.



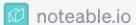
noteable.io

Used Noteable



The introduction to Docker has been added to the notebook. You can view it [here](#).

Next, let's add the section on installing Docker.



noteable.io

Used Noteable



Next, let's add the section on Docker Compose.

 noteable.io

Used Noteable



SH

continue



The section on Docker Compose has been added to the notebook. You can view it [here](#).



Next, let's add the section on Docker in Data Science.

Used Noteable



The section on Docker in Data Science has been added to the notebook. You can view it [here](#).

Finally, let's add the section on exercises. These exercises will give the bootcamp

Finally, let's add the section on exercises. These exercises will give the bootcamp participants a chance to apply what they've learned.



Used **Noteable**



The exercises section has been added to the notebook. You can view it [here](#).

This concludes the creation of the notebook. It provides an introduction to Docker, its installation, basic concepts, and its use in data science, along with some exercises for hands-on practice. You can share this notebook with the bootcamp participants for their self-study and practice.

Please let me know if there's anything else you need.



→ to **noteable** tab!



:::

... +

Introduction to Docker

Docker is a platform that allows users to automate the deployment, scaling, and management of applications. It uses containerization technology to package an application along with its runtime environment, so it can run uniformly across different systems. Docker is not specifically a data science tool, but it is often used in data science environments to ensure consistent and reproducible results across different machines.



This part is in dire need of a fish joke!

Doubleclick or Enter to Edit

+

Installing Docker

The installation process for Docker varies depending on the operating system. Here are the general steps for each major OS:

Windows

1. Visit Docker Hub at <https://hub.docker.com/editions/community/docker-ce-desktop-windows/> and download the Docker Desktop installer.
2. Run the installer, which will guide you through the setup process.

→ to GPT tab!

SH

Rework

<https://app.noteable.io/f/5d287cd0-f6fa-4555-9328-8356c617aa9d?cellID=2baf2527-8494-4dde-9c73-27b0106637ba>

Add a joke with a clownfish meeting the docker whale!

:)

Save & Submit

Cancel



Used Noteable



REQUEST TO NOTEABLE

```
{  
  "file_id": "5d287cd0-f6fa-4555-9328-8356c617aa9d",  
  "cell_id": "2baf2527-8494-4dde-9c73-27b0106637ba"  
}
```



RESPONSE FROM NOTEABLE

```
{  
  "cell": {  
    "id": "2baf2527-8494-4dde-9c73-27b0106637ba",  
    "source": "# Introduction to Docker\n\nDocker is a platform that allows users to  
..."  
  }  
}
```



→ to **noteable** tab!

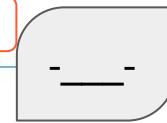
... +

Introduction to Docker

Docker is a platform that allows users to automate the deployment, scaling, and management of applications. It uses containerization technology to package an application along with its runtime environment, so it can run uniformly across different systems. Docker is not specifically a data science tool, but it is often used in data science environments to ensure consistent and reproducible results across different machines.

And now, for a little humor: Why don't clownfish like Docker? Because they're afraid of whales! 🐠🐳

Doubleclick or Enter to Edit



Installing Docker

The installation process for Docker varies depending on the operating system. Here are the general steps for each major OS:

Windows

1. Visit Docker Hub at <https://hub.docker.com/editions/community/docker-ce-desktop-windows/> and download the Docker Desktop installer.
2. Run the installer, which will guide you through the setup process.



My First Project

...



Docker_for_Data_Science



Imputing_Missing_Data



Regular_Expressions_Tutorial



Titanic_0



Titanic



What can you do with a notebook

... +

Markdown ▾



Introduction to Docker

Docker is a platform that allows users to automate the deployment, scaling, and containerization technology to package an application along with its runtime across different systems. Docker is not specifically a data science tool, but it environments to ensure consistent and reproducible results across different environments.

And now, for a little humor: Why don't clownfish like Docker? Because they're

Doubleclick or Enter to Edit

Installing Docker

The installation process for Docker varies depending on the operating system major OS:

Windows

1. Visit Docker Hub at <https://hub.docker.com/editions/community/docker-ce> and download the Docker Desktop installer.
2. Run the installer, which will guide you through the setup process.

macOS

1. Visit Docker Hub at <https://hub.docker.com/editions/community/docker-ce> and download the Docker Desktop installer.

The screenshot shows a Jupyter Notebook interface with a context menu open over a cell. The menu options include: New Notebook..., Import..., Open Project, Move..., Duplicate..., Create Named Version... (with keyboard shortcut cmd + s), Download (.ipynb) (highlighted with a red arrow), Export Notebook (highlighted with a red arrow), and Delete.

Markdown

Introduction to Docker

Docker is a platform that allows users to automate the deployment, scaling, and containerization technology to package an application along with its runtime across different systems. Docker is not specifically a data science tool, but it environments to ensure consistent and reproducible results across different r

And now, for a little humor: Why don't clownfish like Docker? Because they're

Doubleclick or Enter to Edit

Installing Docker

The installation process for Docker varies depending on the operating system major OS:

Windows

1. Visit Docker Hub at <https://hub.docker.com/editions/community/docker-ce> Docker Desktop installer.
2. Run the installer, which will guide you through the setup process.

macOS

1. Visit Docker Hub at <https://hub.docker.com/editions/community/docker-ce>



Table of Contents



- **Introduction to Docker**

**Installing Docker****Windows****macOS****Linux****Docker Basics****Docker Images****Docker Containers****Docker Commands****Building a Docker Image****Dockerfile****Building an Image****Docker Compose****Docker Compose YAML file****Running multi-container Dock...**

... +

Markdown ▾ ?

Introduction to Docker

Docker is a platform that allows users to automate the deployment, scaling, and management of applications by using containerization technology to package an application along with its runtime environment into a single, portable unit. Docker is not specifically a data science tool, but it provides environments to ensure consistent and reproducible results across different environments.

And now, for a little humor: Why don't clownfish like Docker? Because they're Docker!

Doubleclick or Enter to Edit

Installing Docker

The installation process for Docker varies depending on the operating system you are using. Here are the major OS:

Windows

1. Visit Docker Hub at <https://hub.docker.com/editions/community/docker-ce-desktop> and download the Docker Desktop installer.
2. Run the installer, which will guide you through the setup process.

macOS

1. Visit Docker Hub at <https://hub.docker.com/editions/community/docker-ce-desktop> and download the Docker Desktop installer.

And now, for a little humor: Why don't clownfish like Docker? Because they're afraid of whales! 🐠🐳

Markdown ▾ ?

Installing Docker

The installation process for Docker varies depending on the operating system. Here are the general steps for each major OS:

Windows

1. Visit Docker Hub at <https://hub.docker.com/editions/community/docker-ce-desktop-windows/> and download the Docker Desktop installer.
2. Run the installer, which will guide you through the setup process.

macOS

1. Visit Docker Hub at <https://hub.docker.com/editions/community/docker-ce-desktop-mac/> and download the Docker Desktop installer.
2. Run the installer, which will guide you through the setup process.

Linux

The installation process for Linux depends on the distribution. Docker provides a comprehensive set of installation guides for different distributions at <https://docs.docker.com/engine/install/>.

After installation, you can verify that Docker is installed correctly by opening a terminal and running the command



Agenda

- Introduction
- Prompt Engineering Basics
- GPT-4 & Plugin Workflows
- **Prompting with OpenAI and LangChain**
- Evaluating Model Response
- Summary
- Q&A

Prompt Engineering Techniques

Clear Instructions

Context Framing

System Messages

Iterative Refinement

Prompting with OpenAI

1. Read OpenAI API key
2. Model - GPT-3.5-turbo (ChatGPT)
3. Read a jupyter notebook consisting of tasks to be completed
4. Feed the notebook (json format) to the model with the prompt

Listed tasks in a markdown cell

Jupyter Notebook

Introduction

Your task is to

- Read the csv file into a pandas dataframe.
- Find the list of unique cities where matches were played
- Find the columns which contains null values if any ?
- List down top 5 most played venues
- Get top 5 goal scorers of the tournament

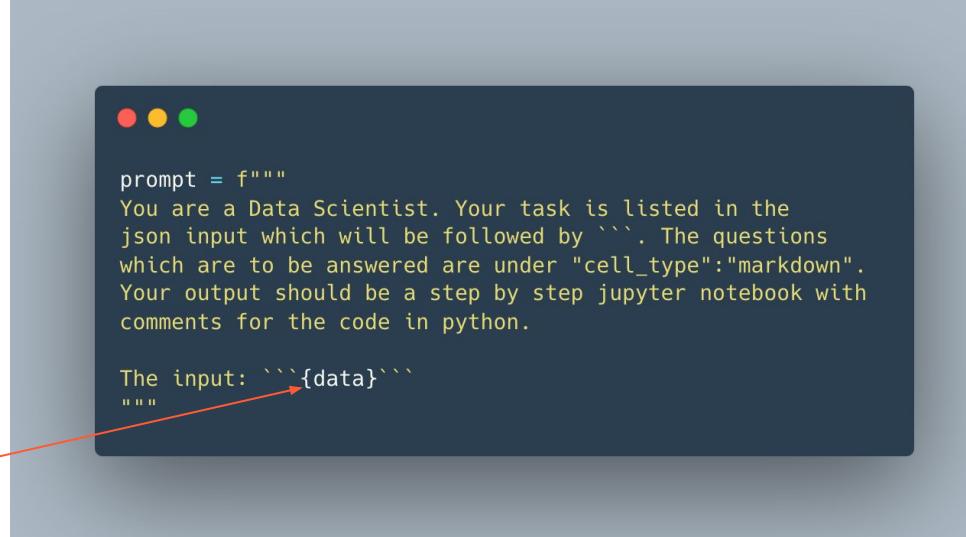
Prompting with OpenAI

We create a prompt

- consists of a task to be followed.
- tells the model what input questions it has to work on

The input file, which lists the tasks

Prompt #1



```
prompt = f"""
You are a Data Scientist. Your task is listed in the
json input which will be followed by ``` . The questions
which are to be answered are under "cell_type": "markdown".
Your output should be a step by step jupyter notebook with
comments for the code in python.

The input: ``{data}````
```

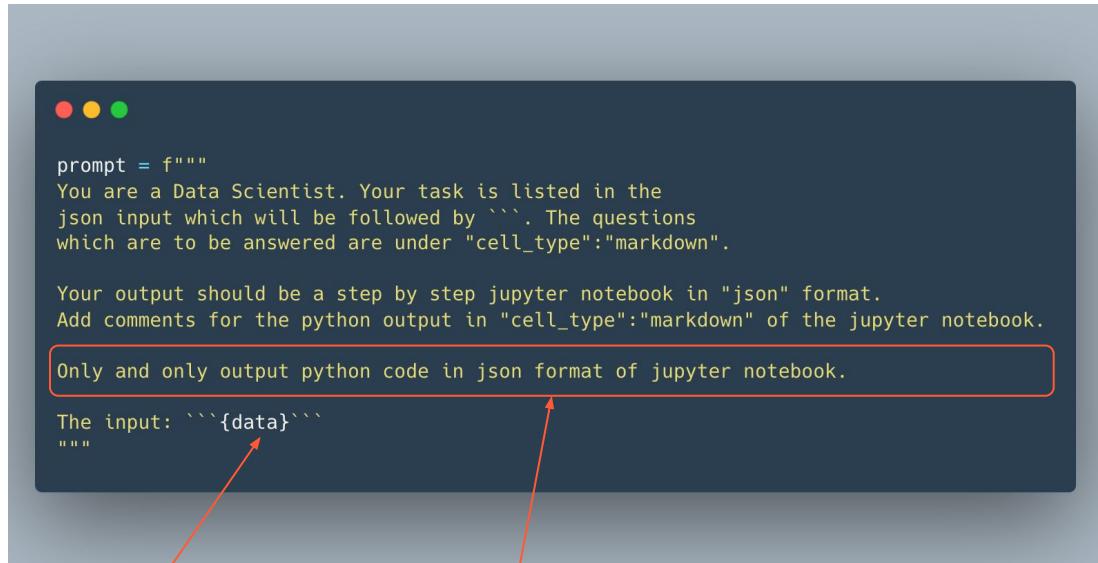
A screenshot of a terminal window showing a Python script. The script defines a variable 'prompt' as a multi-line string. It instructs the AI to act as a Data Scientist, to follow tasks listed in json input, and to respond with a step-by-step jupyter notebook. It also specifies that the AI should use triple quotes for code blocks and provide comments for the code. A red arrow points from the text 'The input: ``{data}````' to the placeholder '{data}' in the string.

Prompting with OpenAI

Following iterative refinement

- Clear instructions
- Restricting the thinking of the model
- Forcing it to only output what we need

Prompt #2



The screenshot shows a dark-themed AI interface window. At the top, there are three colored dots (red, yellow, green). Below them, the code for the prompt is shown:

```
prompt = f"""
You are a Data Scientist. Your task is listed in the
json input which will be followed by ```. The questions
which are to be answered are under "cell_type":"markdown".

Your output should be a step by step jupyter notebook in "json" format.
Add comments for the python output in "cell_type":"markdown" of the jupyter notebook.

Only and only output python code in json format of jupyter notebook.

The input: ```{data}```
"""

```

Two red arrows point from the text "The input: ```{data}```" and the word "Refinement" at the bottom left to the corresponding parts in the code block above.

The input file, which lists the tasks

Refinement

Prompting with LangChain



What is LangChain?

Framework designed to simplify the creation of applications using large language models (LLMs)

Why LangChain?

- gives flexibility to use different prompt templates.
- templates can be used for Few-Shot learning.

Prompting with LangChain



Prompt Templates

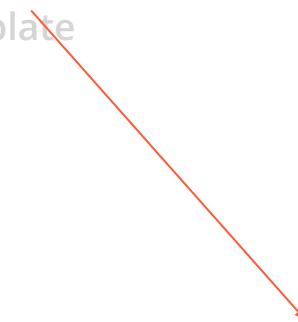
- System Message Prompt Template
- Human Message Prompt Template
- AI Message Prompt Template
- Chat Prompt Template

Prompting with LangChain



Prompt Templates

- **System Message Prompt Template**
- Human Message Prompt Template
- AI Message Prompt Template
- Chat Prompt Template



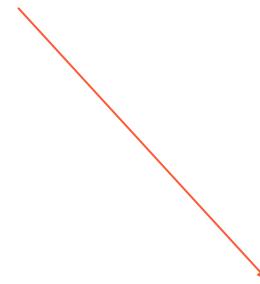
Provides a default message to the model, saying what it's actual task is.

Prompting with LangChain



Prompt Templates

- System Message Prompt Template
- Human Message Prompt Template
- AI Message Prompt Template
- Chat Prompt Template



Gives the model an example prompt, which it can expect from the users.

Prompting with LangChain



Prompt Templates

- System Message Prompt Template
- Human Message Prompt Template
- **AI Message Prompt Template**
- Chat Prompt Template



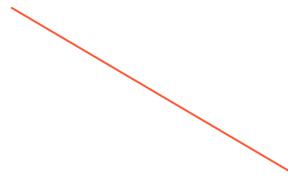
Tells the model what kind of response it should give to the user's question.

Prompting with LangChain



Prompt Templates

- System Message Prompt Template
- Human Message Prompt Template
- AI Message Prompt Template
- Chat Prompt Template



Combines all the Prompt Templates together.

Prompting with LangChain

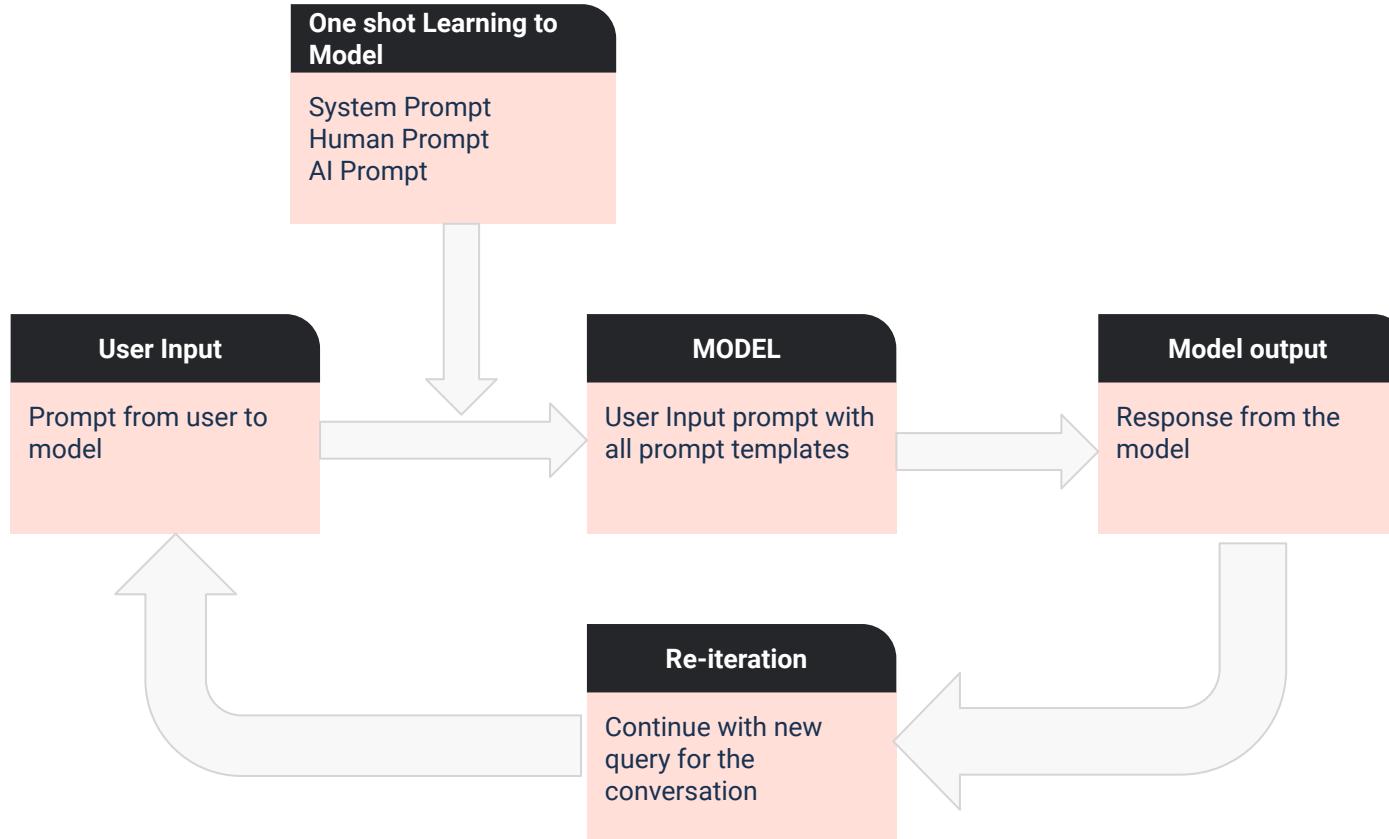


Prompt Templates

- System Message Prompt Template
- Human Message Prompt Template
- AI Message Prompt Template
- Chat Prompt Template

- Keeps track of previous conversations
- Can keep limited memory or summary
- No need to input whole prompt, like in OpenAI

Prompting with LangChain



Prompting with LangChain



Use Cases

- Generate better responses in few iterations
- Generate solutions for the tasks
- Generate tasks notebooks on a specific topic

Agenda

- Introduction
- Prompt Engineering Basics
- GPT-4 & Plugin Workflows
- Prompting with OpenAI and LangChain
- Evaluating Model Response
- Summary
- Q&A

Evaluating Model Responses

Evaluating responses using Rubric

Evaluating Model Responses

Evaluating responses using **Rubric**



is typically an evaluation tool or set of guidelines, instructions or rules.

Evaluating Model Responses

Evaluating responses using Rubric

- Model response goes to another LLM
- New LLM model has another set of instructions
- New LLM model
 - is given ideal prompt and response scenario
 - compares response from model #1 with ideal response
 - is given set of choices to evaluate the response on

Evaluating Model Responses

System prompt to model for evaluation



Compare the factual content of the submitted answer with the expert answer. DO NOT ignore any differences in style or punctuation.

The submitted answer may either be a subset or superset of the expert answer, or it may conflict with it. Determine which case applies. Answer the question by selecting one of the following options:

- (A) The submitted answer is a part of the expert answer and is fully consistent with it.
- (B) The submitted answer contains all the same details as the expert answer.
- (C) There is a disagreement between the submitted answer and the expert answer.
- (D) The answers differ, but these differences don't matter from the perspective of factuality.

Evaluating Model Responses

Tells model to compare the response



Compare the factual content of the submitted answer with the expert answer. DO NOT ignore any differences in style or punctuation.

The submitted answer may either be a subset or superset of the expert answer, or it may conflict with it. Determine which case applies. Answer the question by selecting one of the following options:

- (A) The submitted answer is a part of the expert answer and is fully consistent with it.
- (B) The submitted answer contains all the same details as the expert answer.
- (C) There is a disagreement between the submitted answer and the expert answer.
- (D) The answers differ, but these differences don't matter from the perspective of factuality.

Evaluating Model Responses

Specifying exactly what to focus on



Compare the factual content of the submitted answer with the expert answer. **DO NOT ignore** any differences in style or punctuation.

The submitted answer may either be a subset or superset of the expert answer, or it may conflict with it. Determine which case applies. Answer the question by selecting one of the following options:

- (A) The submitted answer is a part of the expert answer and is fully consistent with it.
- (B) The submitted answer contains all the same details as the expert answer.
- (C) There is a disagreement between the submitted answer and the expert answer.
- (D) The answers differ, but these differences don't matter from the perspective of factuality.

Evaluating Model Responses

Specifying what the response can be if not exact to ideal



Compare the factual content of the submitted answer with the expert answer. DO NOT ignore any differences in style or punctuation.

The submitted answer may either be a subset or superset of the expert answer, or it may conflict with it. Determine which case applies. Answer the question by selecting one of the following options:

- (A) The submitted answer is a part of the expert answer and is fully consistent with it.
- (B) The submitted answer contains all the same details as the expert answer.
- (C) There is a disagreement between the submitted answer and the expert answer.
- (D) The answers differ, but these differences don't matter from the perspective of factuality.

Evaluating Model Responses

Tells model what choices it has, to make a decision



Compare the factual content of the submitted answer with the expert answer. DO NOT ignore any differences in style or punctuation.

The submitted answer may either be a subset or superset of the expert answer, or it may conflict with it. Determine which case applies. Answer the question by selecting one of the following options:

- (A) The submitted answer is a part of the expert answer and is fully consistent with it.
- (B) The submitted answer contains all the same details as the expert answer.
- (C) There is a disagreement between the submitted answer and the expert answer.
- (D) The answers differ, but these differences don't matter from the perspective of factuality.

Agenda

- Introduction
- Prompt Engineering Basics
- GPT-4 & Plugin Workflows
- Prompting with OpenAI and LangChain
- Evaluating Model Response
- **Summary**
- Q&A

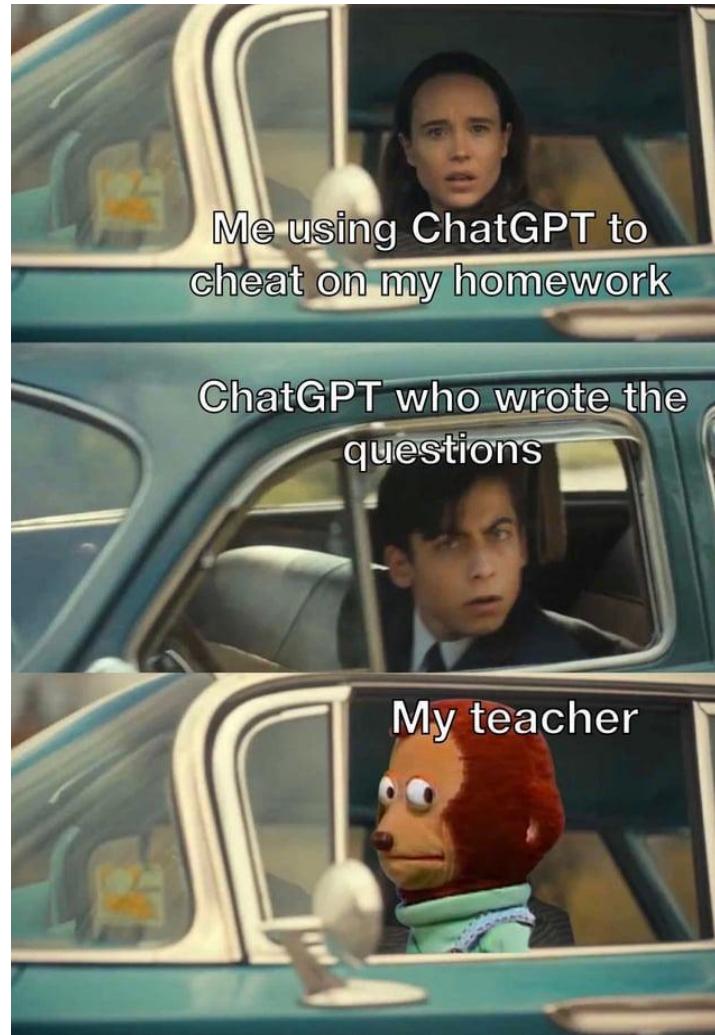
Recap

- Design and leverage your environment, one mini-habit at a time.
Go for low hanging fruits, allocate time for experimentation
- Iterative prompting with OpenAI
- Conversation with LLM using LangChain
 - Creating solutions for tasks
 - Creating tasks based on a topic
- Evaluating model response
 - check model responses before showing the user

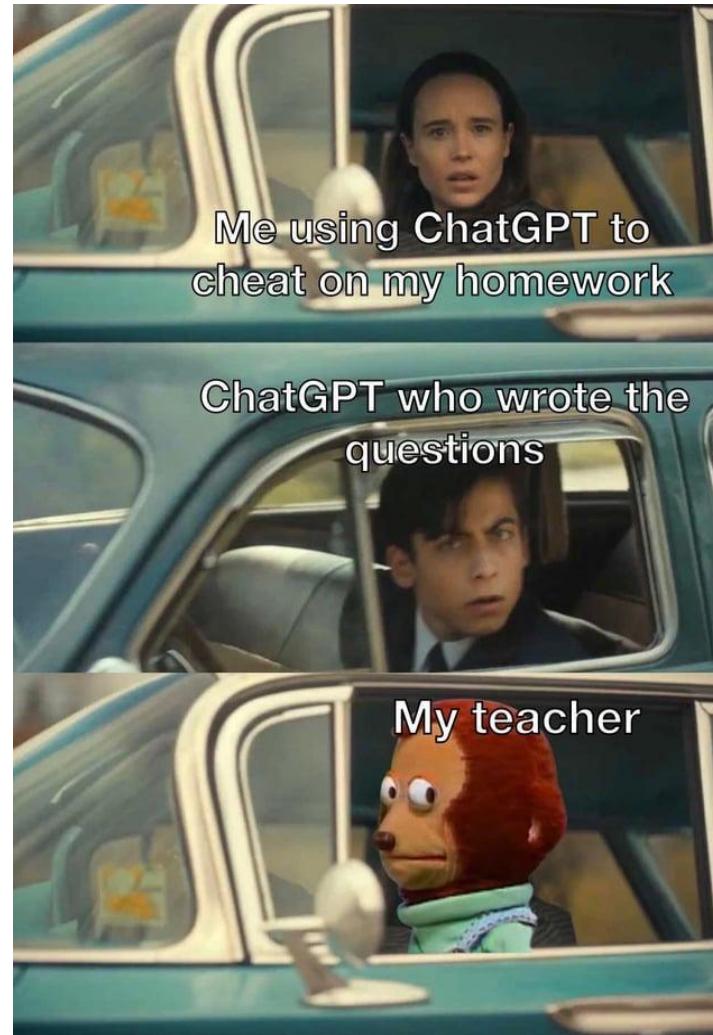
From Brute Force to Finesse

Clarity is key.

**Know what you want in well-defined
words**



**Who knows.
Someday,
there will be
just AIs talking
to each other.**





https://commons.wikimedia.org/wiki/File:Die_Grenzen_meiner_Sprache_bedeuten_die_Grenzen_meiner_Welt._Ludwig_Wittgenstein,_1889-1951_-de.png

The
limits of
my language
mean the
limits of
my world

Thank you for the attention.
Go explore! Our world got a lot bigger.



https://commons.wikimedia.org/wiki/File:Die_Grenzen_meiner_Sprache_bedeuten_die_Grenzen_meiner_Welt._Ludwig_Wittgenstein,_1889-1951_-de.png

The
limits of
my language
mean the
limits of
my world

References

- <https://learn.deeplearning.ai/chatgpt-building-system>
- <https://learn.deeplearning.ai/chatgpt-prompt-eng>
- <https://learn.deeplearning.ai/langchain>
- <https://platform.openai.com/overview>
- <https://medium.com/dair-ai/prompt-engineering-lecture-71099d8cbb9e>



Icons

www.vecteezy.com

www.flaticon.com

www.freepik.com



QnA Time

Notebook to try prompting

- We have shared the link with you. Please look into the chat :)



Kontakt

Noch Fragen? So könnt ihr uns erreichen:

Online:

studienberatung@neuefische.de
www.neuefische.de

Offline:

neue fische GmbH
Gasstraße 6a
22761 Hamburg

+49 40 2285 9616 