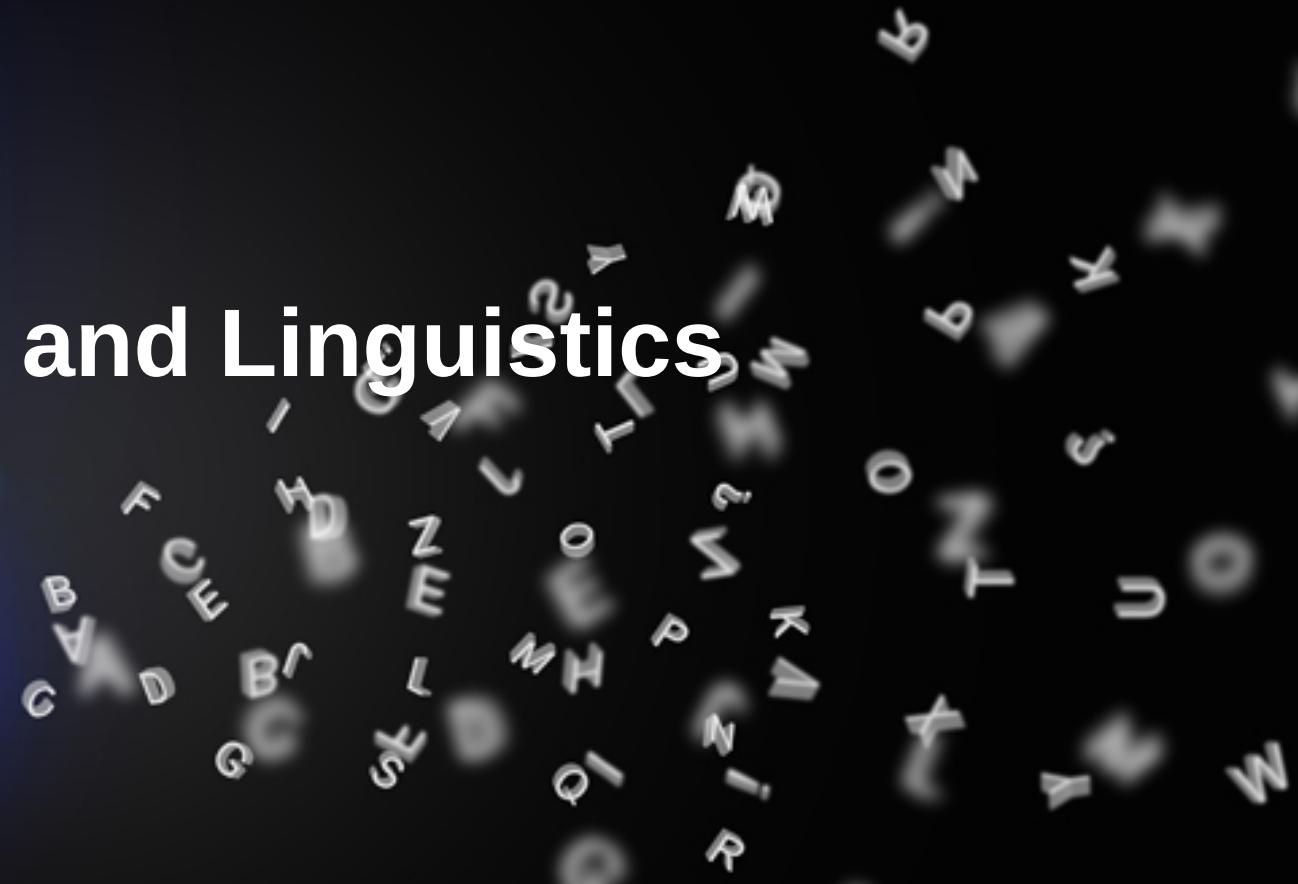


AI, Prompt Engineering, and Linguistics

Speaker: Dr. Haowen Jiang

Date: May 8th, 2024



Outline

- My background
- Intro to AI
- Large Language Models
- Prompt Engineering & Linguistics
- Diffit: an AI tool for teaching
- Q & A

My background



Education



Nat'l Taiwan University

MA, Linguistics
BA, Foreign Languages



Rice University

PhD, Linguistics
MA, Linguistics



Experience



English lecturer

- at Nat'l Taipei U of Technology

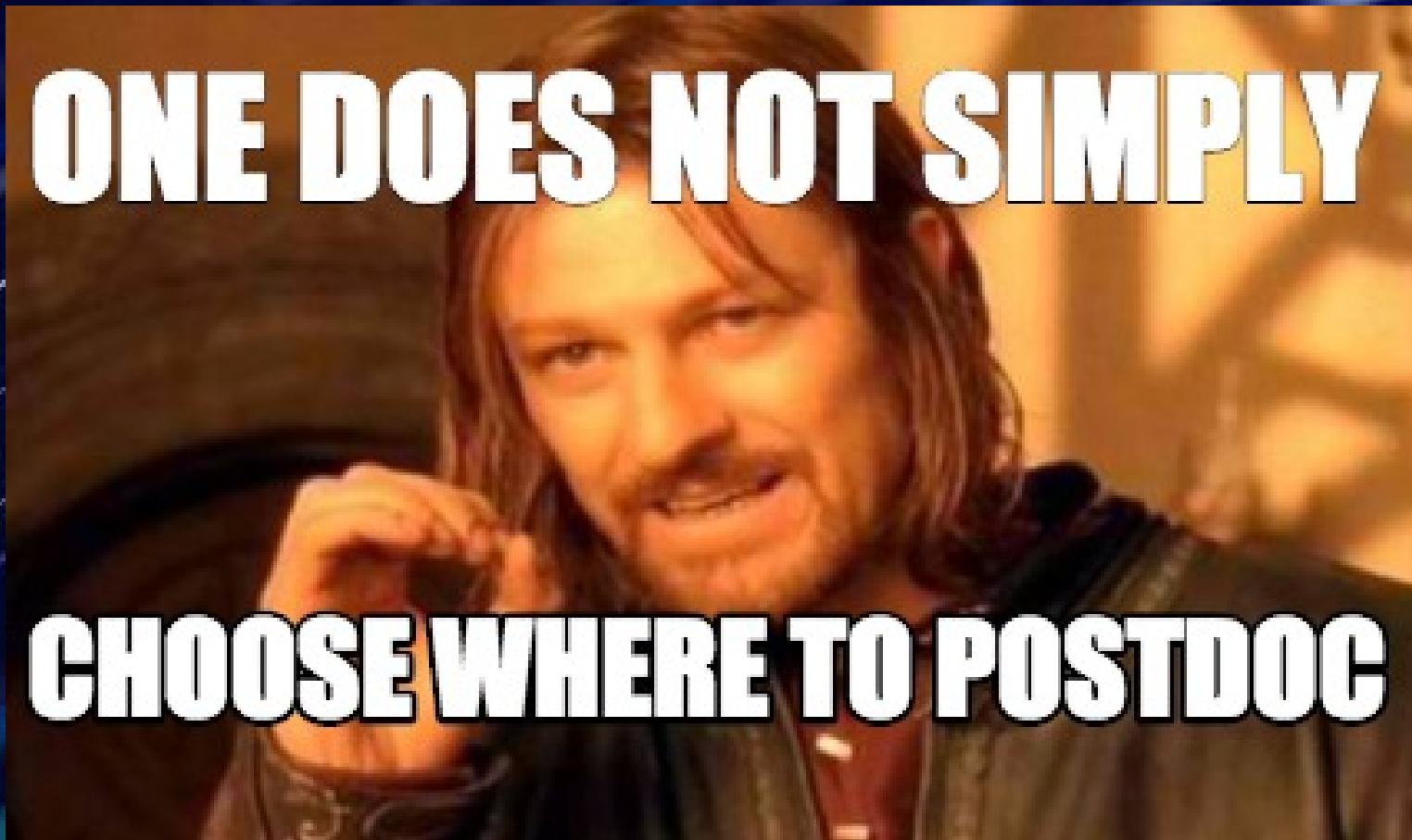
I LOVE GRAMMAR.

IT MAKES PEOPLE CRY.

weknowmemes

Experience

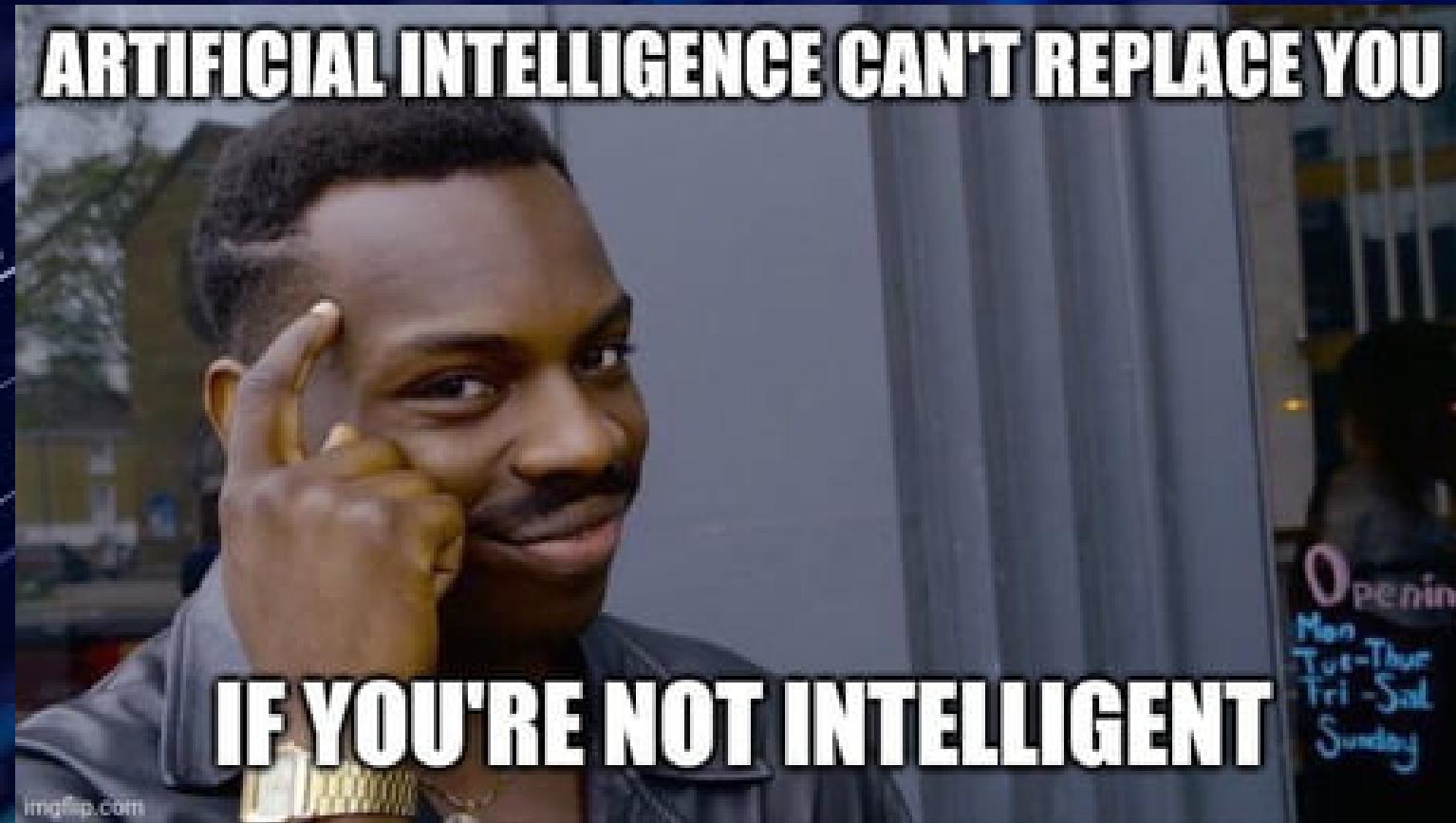
 **Postdoc researcher**
• at Peking U



Experience



- at Hamastar Tech
- at Wisers Information Limited



Experience

 Assistant manager
of data analysis

- at Shin Kong Life

**YOU GOT INSURANCE FOR YOUR
CAR**

**BUT NO LIFE INSURANCE FOR YOU
AND YOUR FAMILY?**

makeameme.org

Intro to AI

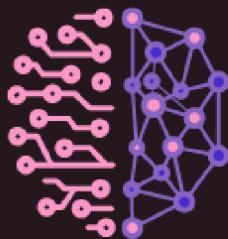


AI and its other names

AI

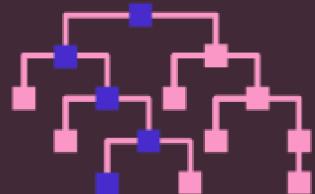
Artificial Intelligence

Any technique that enables computers to mimic human behavior.



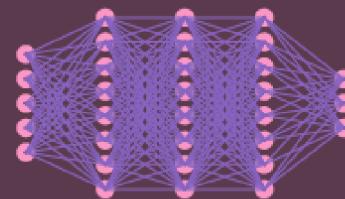
Machine Learning

The ability to learn without directly being programmed.



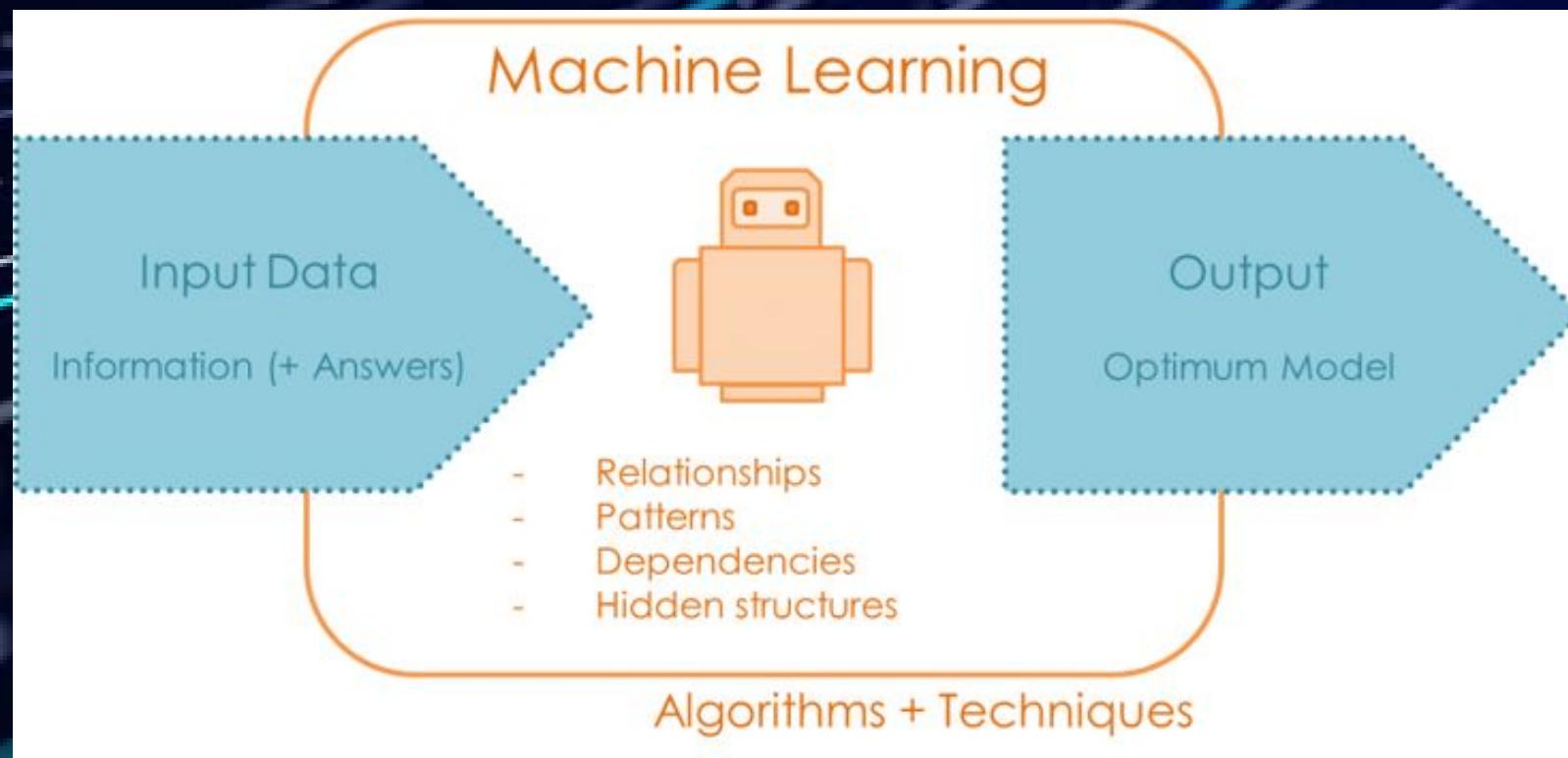
Deep Learning

The learning of underlying features in data using deep neural networks.



Searching for a function

The goal of ML/DL is to search for a *function* that takes some input and then produces some output in a way that humans would normally do.



Function

Linguistics

LIKE(I, languages)

Programming

```
def like(subj, obj):  
    print(f"{subj} like {obj}.")
```

```
like("I", "languages")
```

Try it out [here!](#)

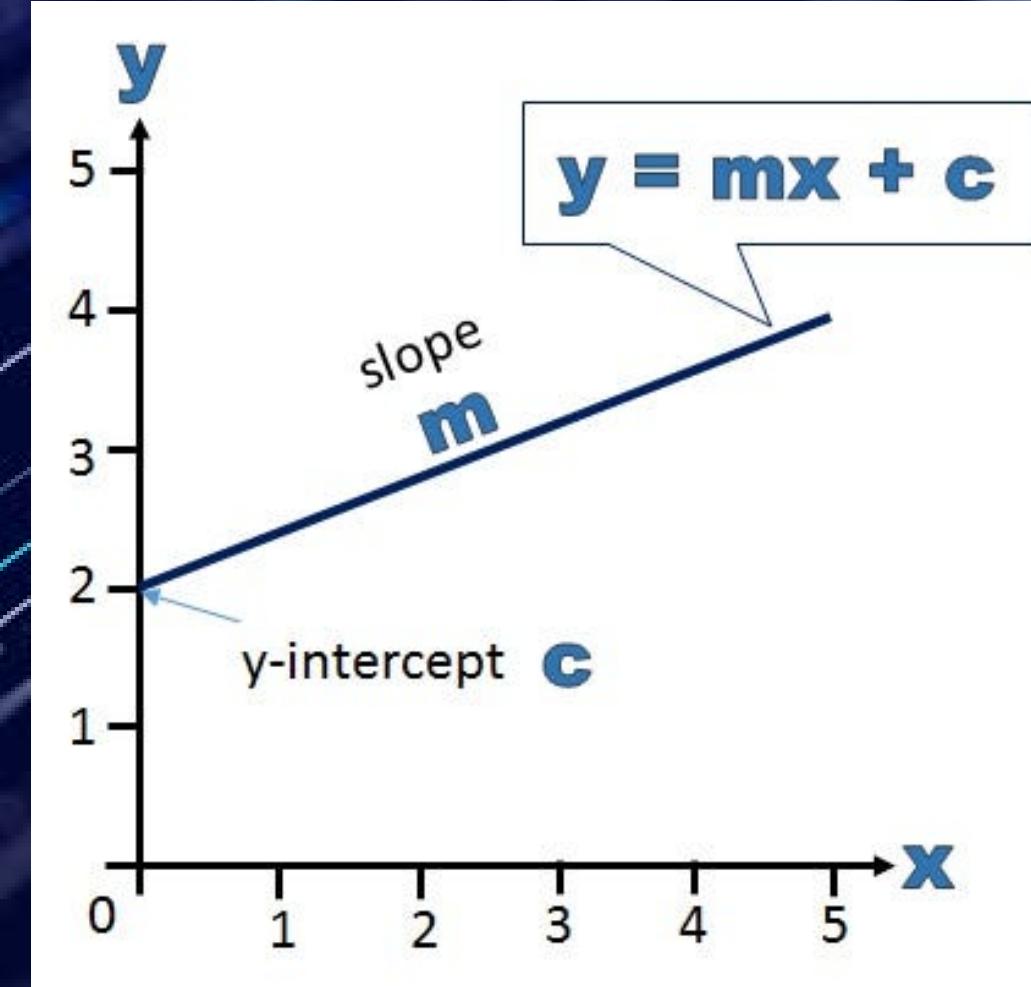
Data-driven AI

A data-driven model
is trained on data points instead of being
coded upfront.

```
inputs = [2, 3, 4] # the x variable  
outputs = [12, 17, 22] # the y variable
```

Random : $y = 0.1x + 5$

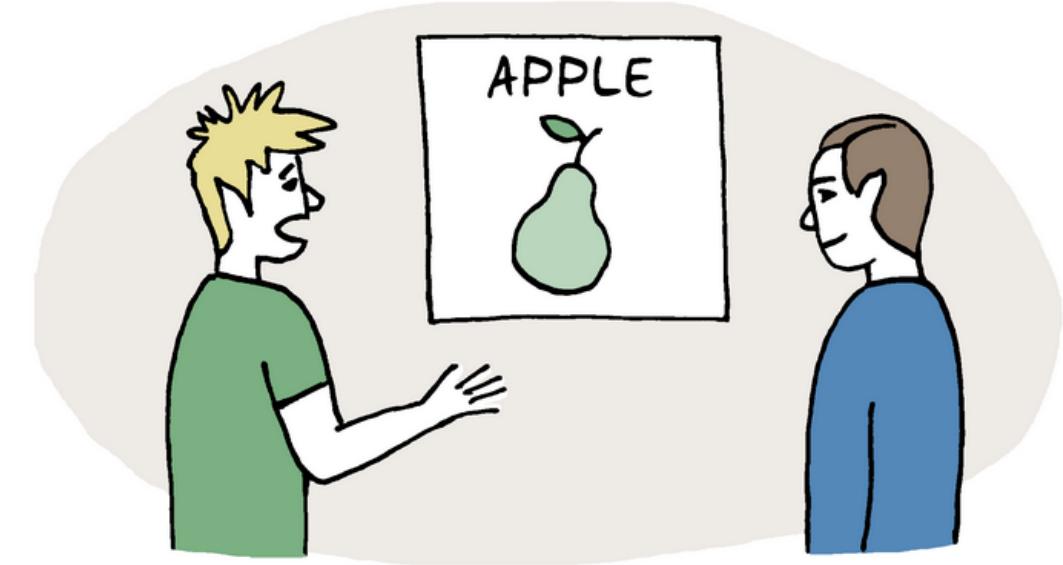
Trained : $y = 5x + 2$



How do machines learn?

Machines learn
by trial and error,
just as humans do.
</div>

MACHINE LEARNING



WELL, A MORE ACCURATE
NAME WOULD BE
MACHINE GUESSING



Dataedo /cartoon

Piotr@Dataedo

Why is machine learning powerful?

True power of ML/DL

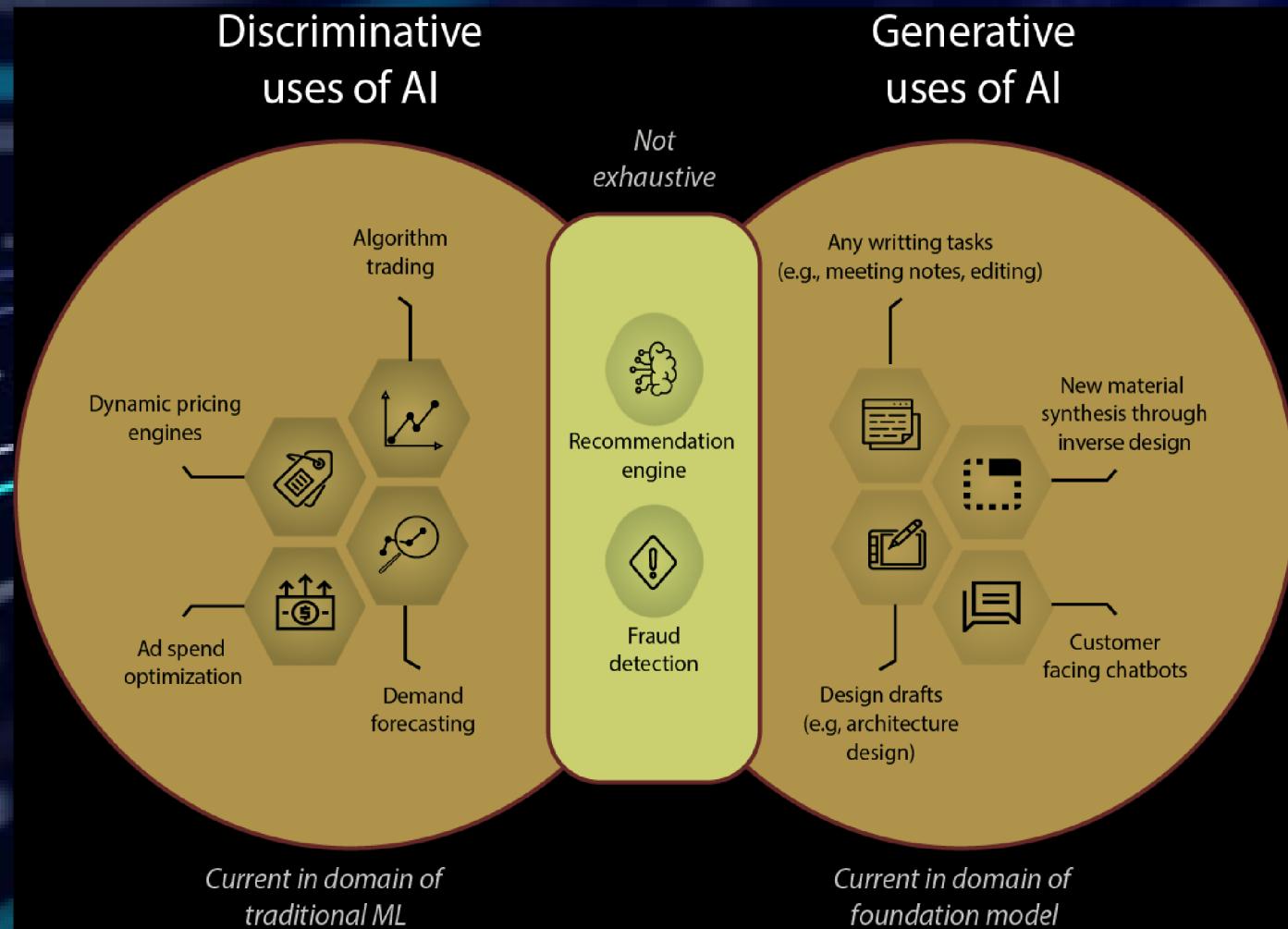
- humans only need to provide input and output
- computers are in charge of figuring out the right process (i.e. a model)

NEW COMPUTER MODEL



SUPPORTED MY HYPOTHESIS
imgflip.com

2 types of AI



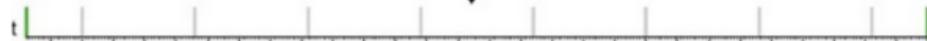
2 types of discriminative models

Regression



What will be the temperature tomorrow?

84°



Fahrenheit

Classification



Will it be hot or cold tomorrow?

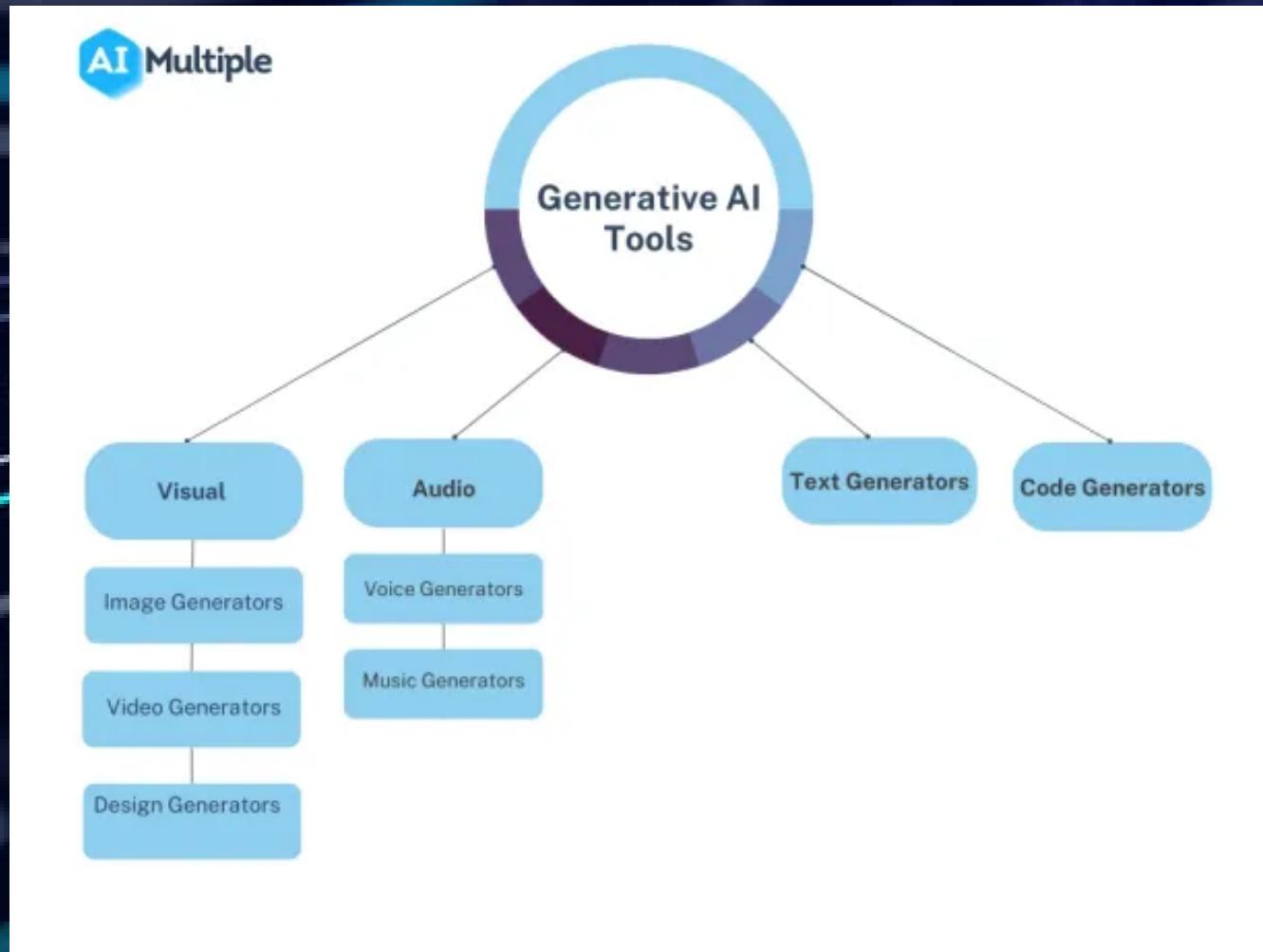
COLD

HOT



Fahrenheit

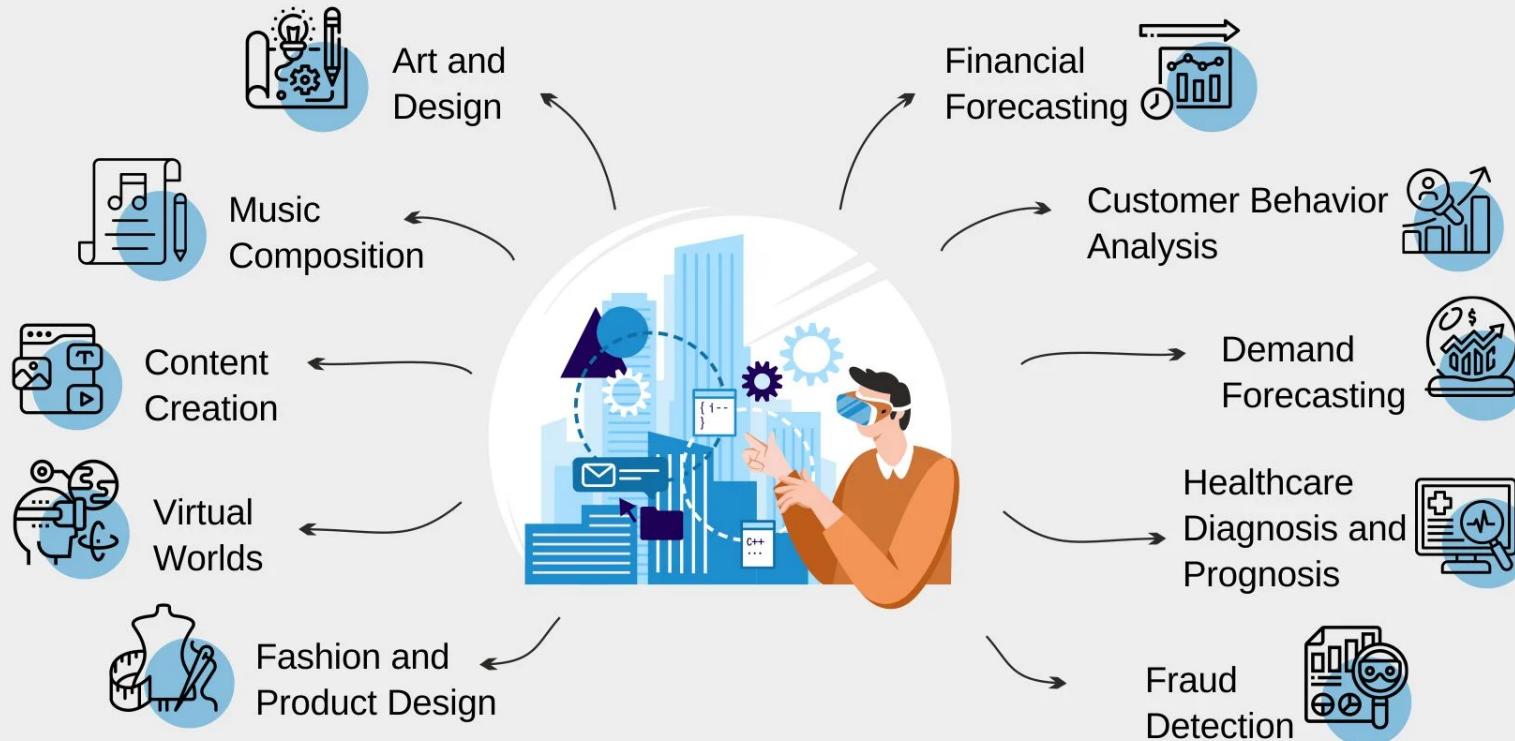
4 types of generative models



Use cases of GenAI

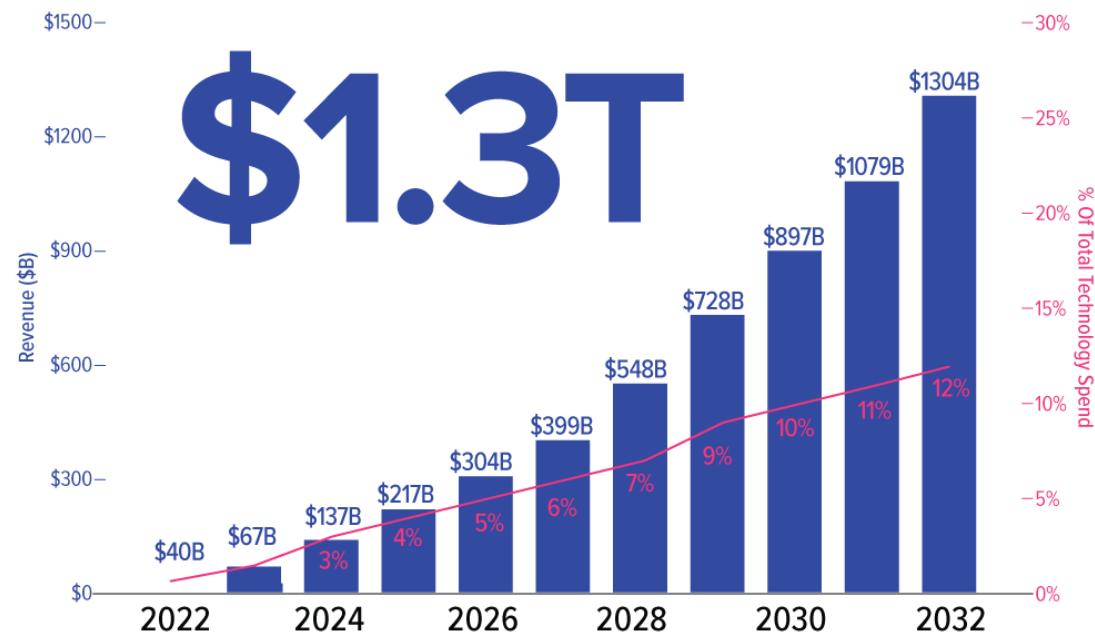
Generative AI Applications

Neebal
Technologies
www.neebal.com



Revenue forecast of GenAI

Generative AI Market Forecast by Revenue and Technology Spend



Source: Bloomberg 2023

Synthetia

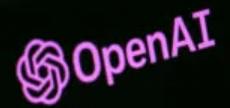
Large Language Models

UNDERSTANDING THEIR IMPACT

Exploring Large
Language Models

ChatGPT's status

The New York Times
The best artificial
intelligence chatbot
ever released to the
general public.



ChatGPT: Optimizing Language Models for Dialogue

We've trained a model called ChatGPT which interacts in a conversational way. The dialogue format makes it possible for ChatGPT to answer followup questions, admit its mistakes, challenge incorrect premises, and reject inappropriate requests. ChatGPT is a sibling model to InstructGPT, which is designed to follow an instruction in a prompt and provide a response.

ChatGPT's significance

TechGoing

ChatGPT's history is as significant as the birth of the PC or the Internet

~ Bill Gates

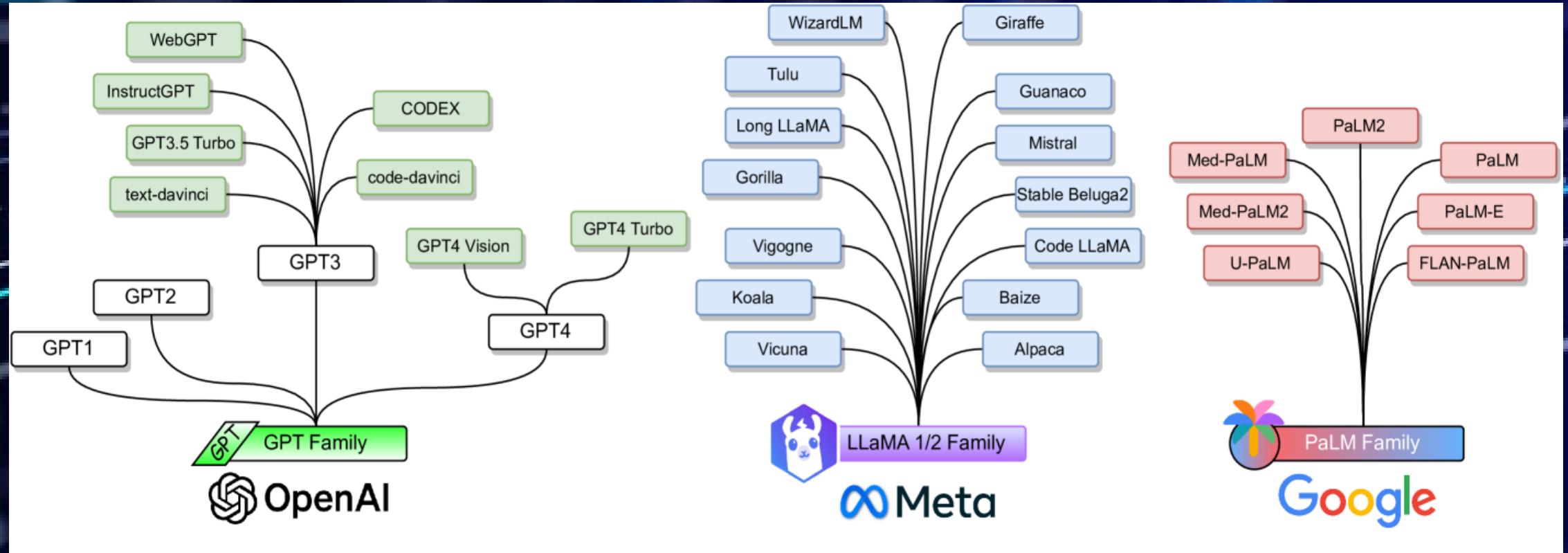


ChatGPT's model

ChatGPT is powered by a large language model (LLM) called Generative Pre-trained Transformer (GPT).



The LLM family



Trends of LLM

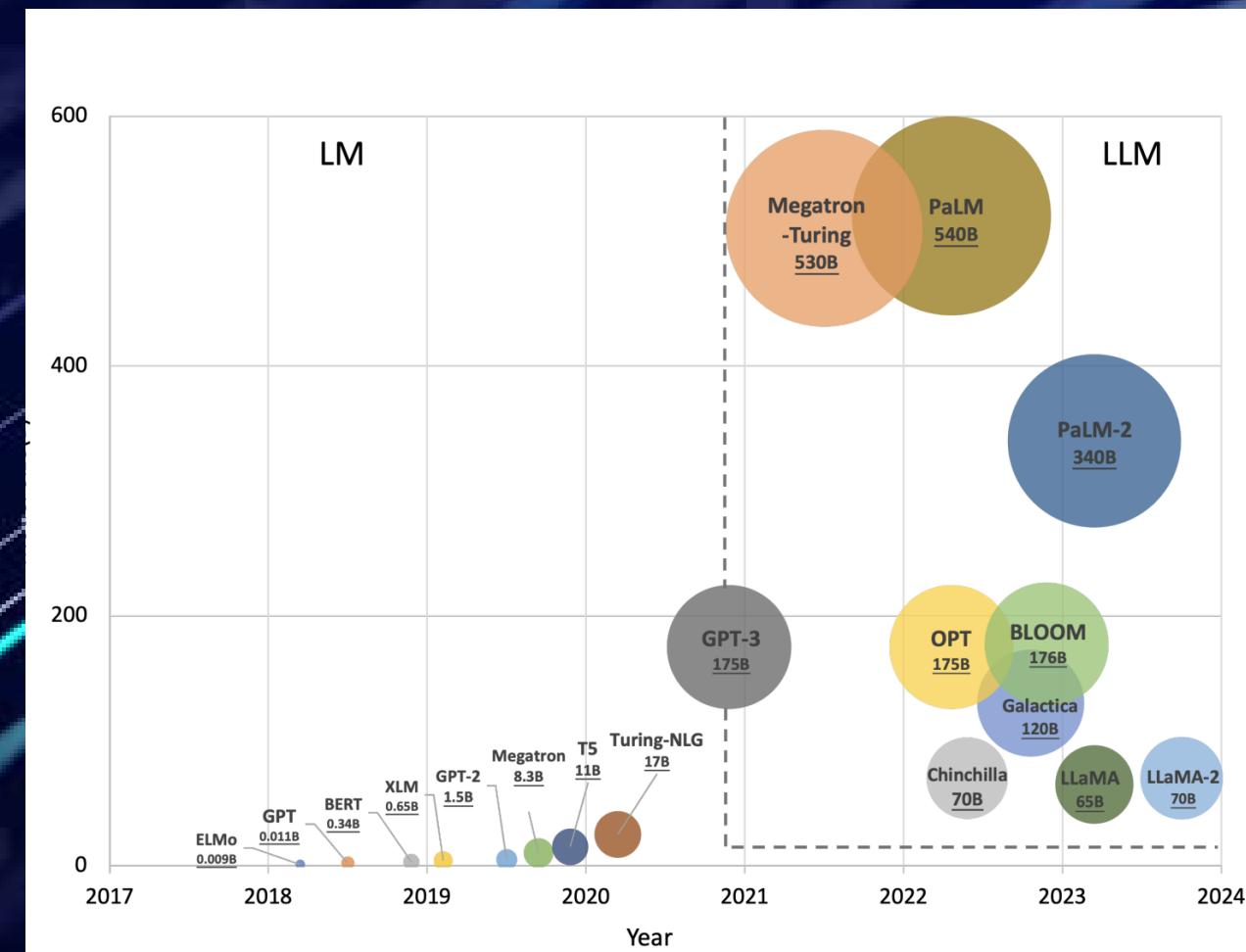
LLM became a popular term in Taiwan around Sep, 2023.



What makes LLM large?

LLM

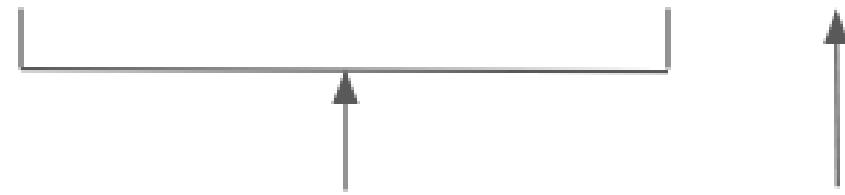
- is large because of the size of its parameters.
- Recall that $y = 0.1x + 5$ has one parameter.



What is a language model?

A language model predicts the next word based on conditional probability.

$S = \text{Where are we going}$

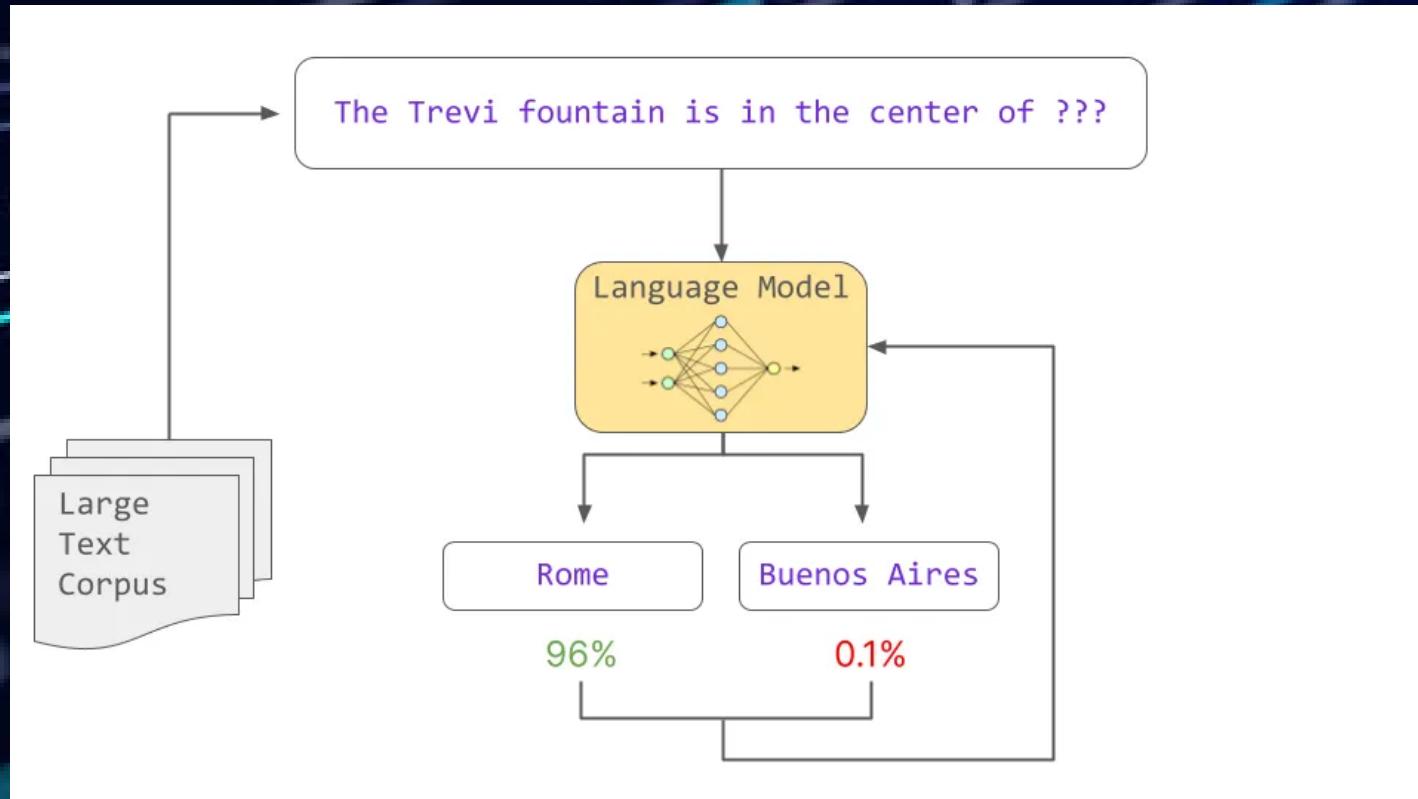


Previous words
(Context) Word being
predicted

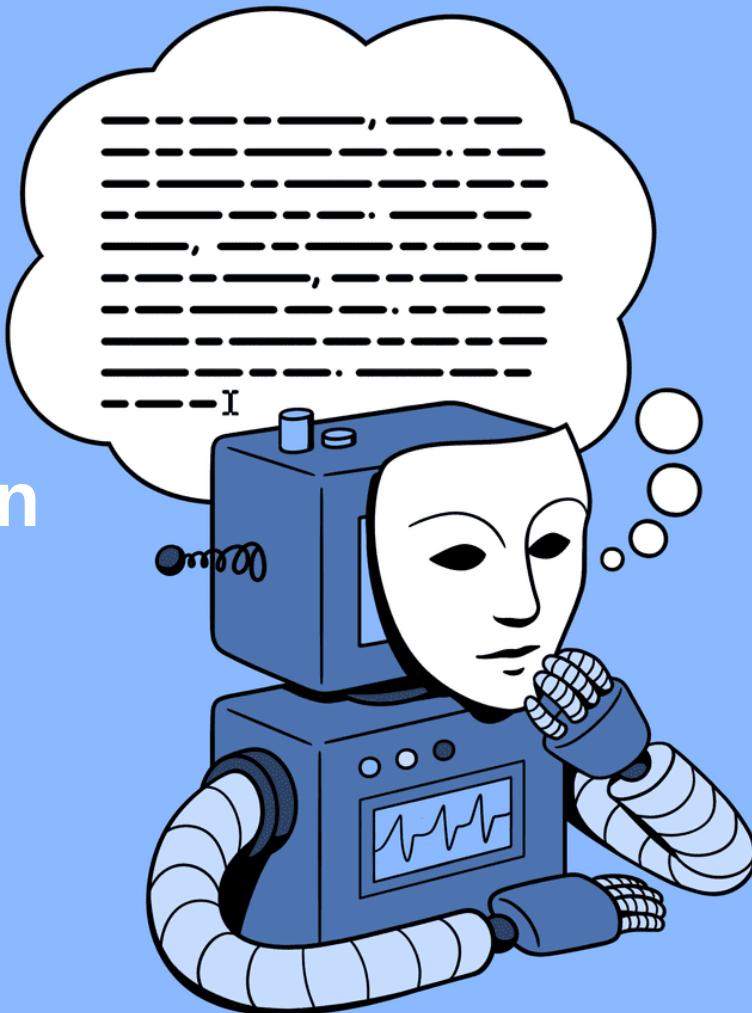
$$P(S) = P(\text{Where}) \times P(\text{are} \mid \text{Where}) \times P(\text{we} \mid \text{Where are}) \times P(\text{going} \mid \text{Where are we})$$

A language model has some world knowledge.

Given a huge corpus of texts, a language model can acquire some basic world knowledge.



Definition



Large Language Model (LLM)

[lärj 'laŋ-gwij 'mä-dəl]

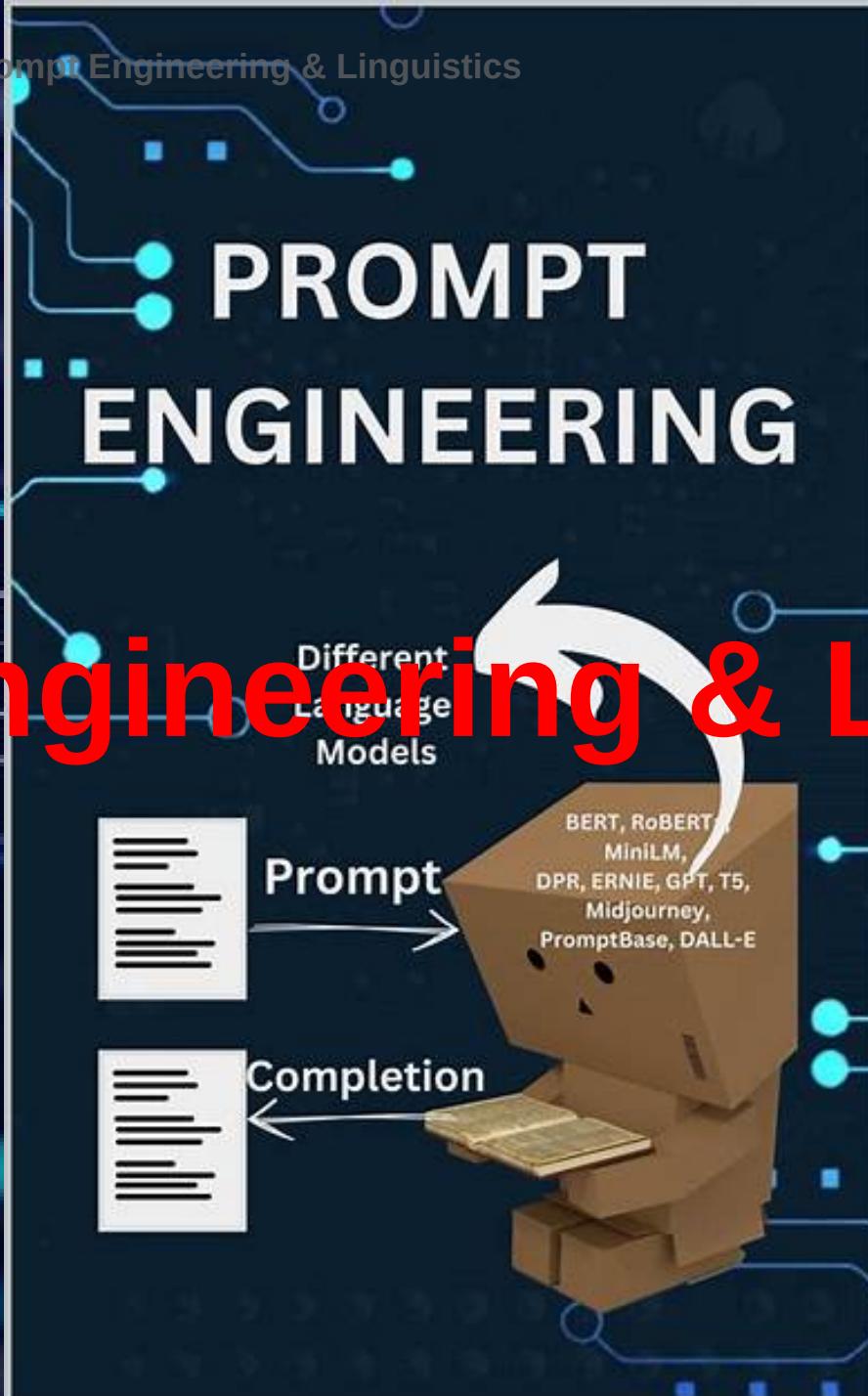
A deep learning algorithm that's equipped to summarize, translate, predict, and generate human-sounding text to convey ideas and concepts.

Top AI chatbots in 2024 powered by LLM

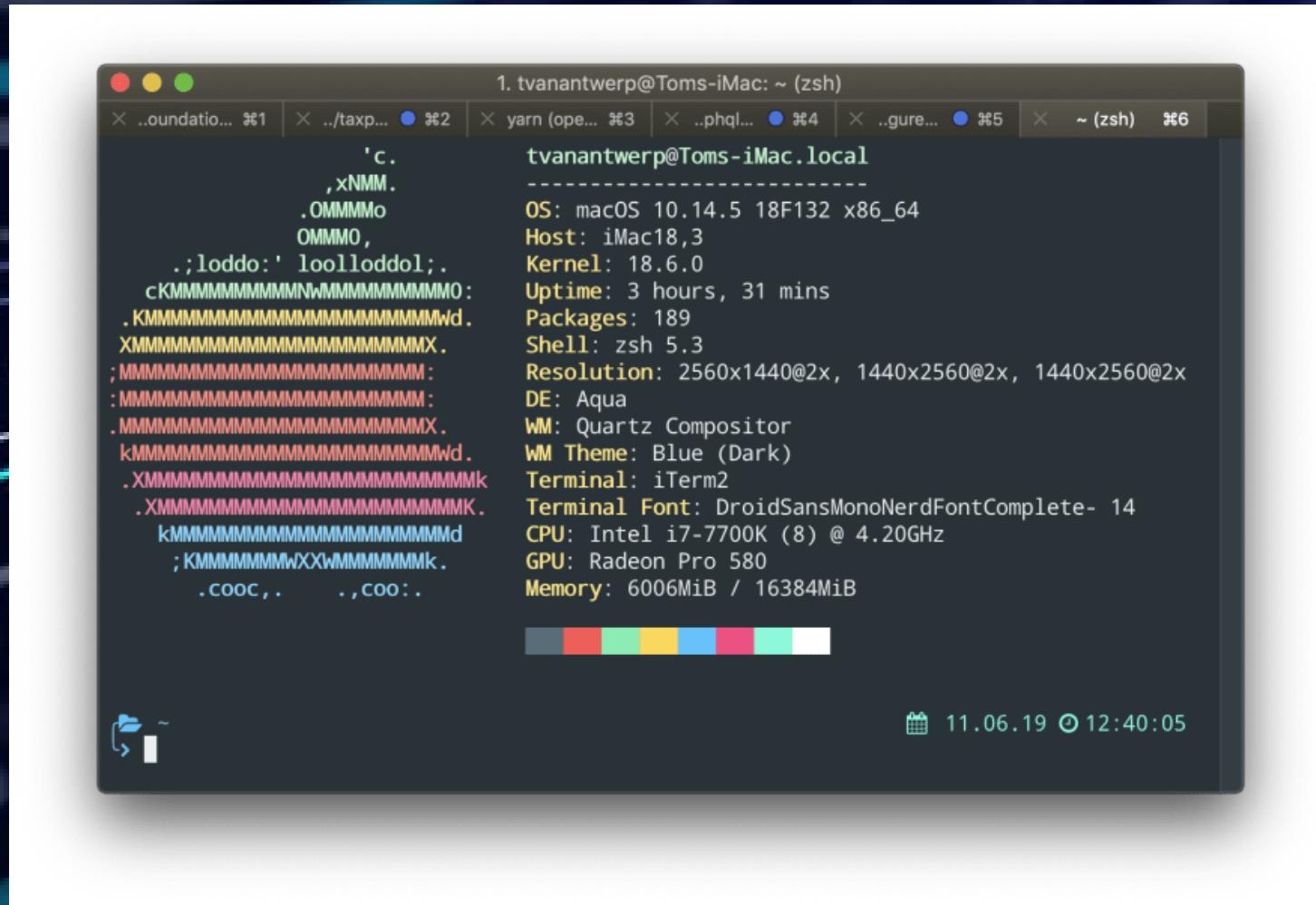
	 ChatGPT	 Google Bard	 Claude AI	 Bing Chat	 OORT AI
Release date	Nov. 2022	Mar. 2023	Mar. 2023	Mar. 2023	Sept. 2023
Key feature	Creativity and human-like conversation	Google ecosystem assistant	Safety and ethics response	Web-based contextual response	Privacy and customization
Language model	GPT-3.5/Turbo	Gemini (successor to LaMDA and PaLM)	Claude	GPT-4	Enhanced vicuna
Information access	Offline knowledge data	Online internet data	Offline knowledge data	Online internet data	Offline knowledge data
Customer data storage	Centralized	Centralized	Centralized	Centralized	Decentralized
Integration API	Available	Not available	Available	Not available	Available

<div class="caption"> Source Article </div>

Prompt Engineering & Linguistics

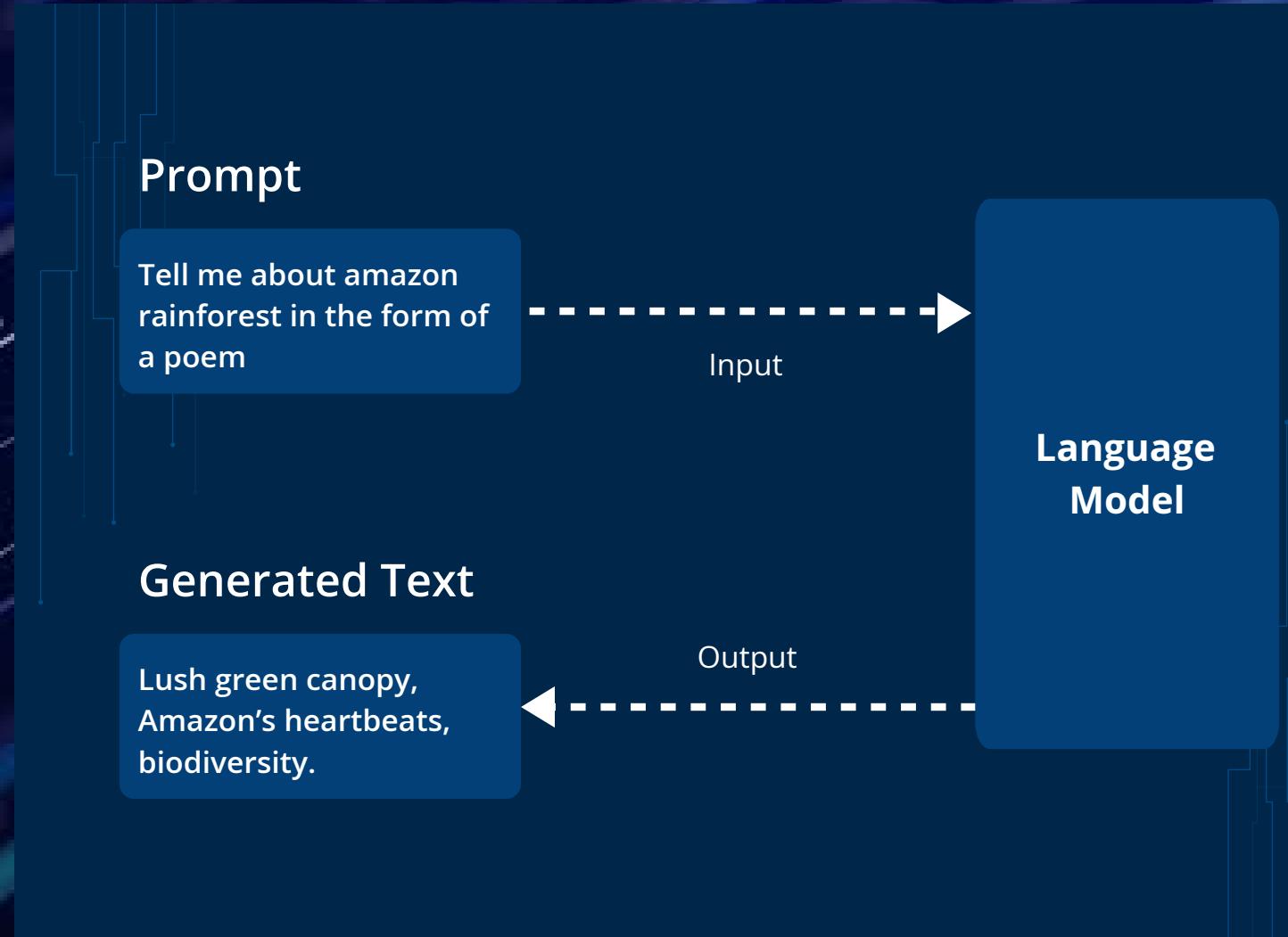


Prompt on the terminal

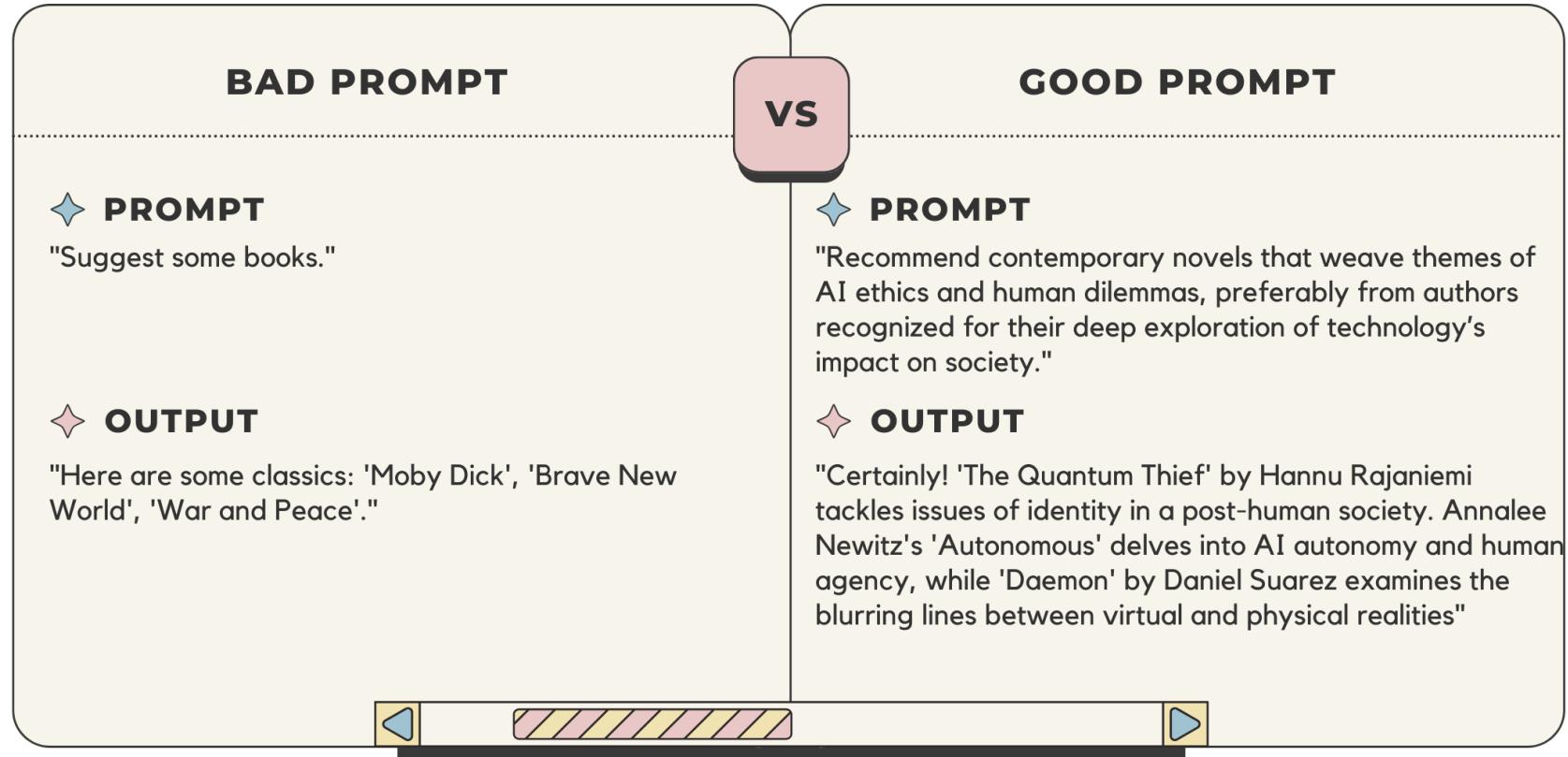


Prompt in GenAI

Prompt
In the context of GenAI, a prompt is a specific instruction or input given to the AI model to generate a desired output.

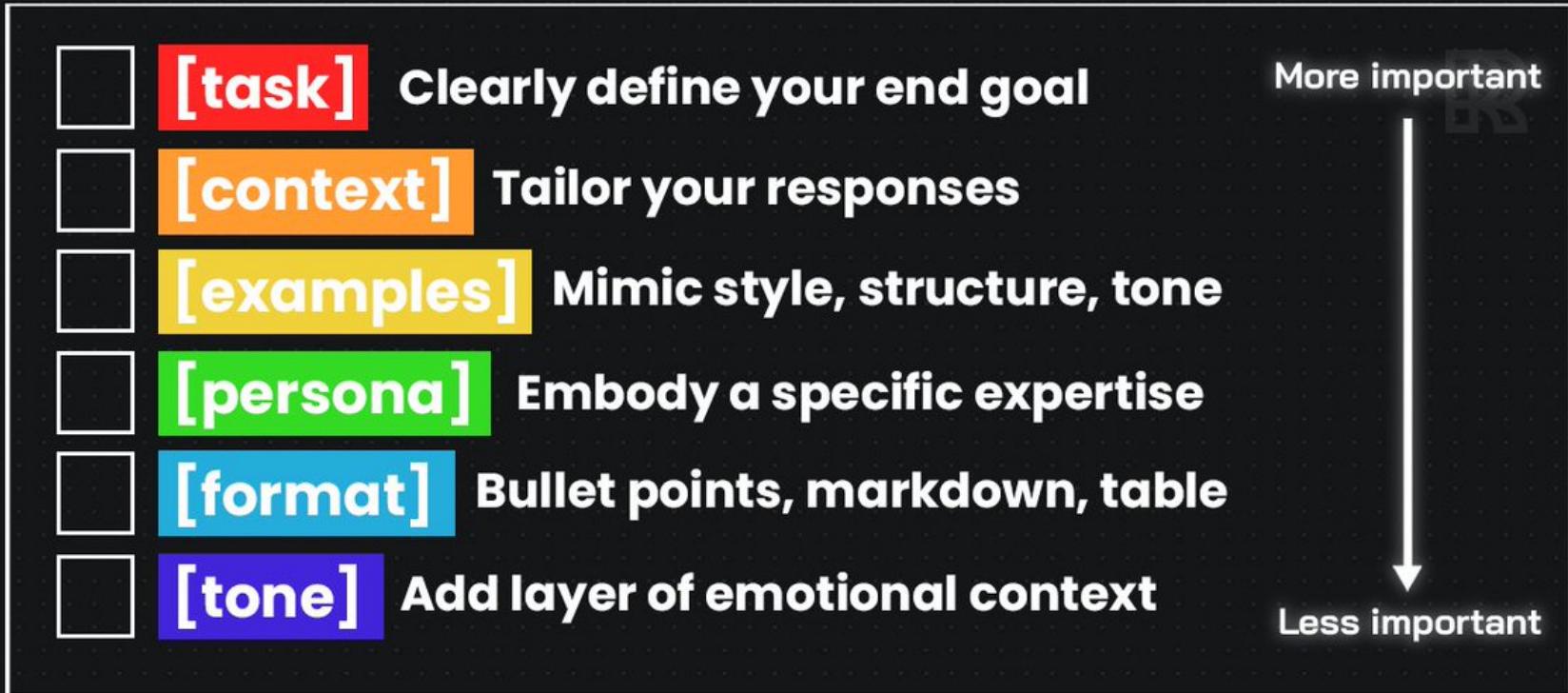


What makes a good prompt?



Elements of a good prompt

The 6-Step Prompt Checklist

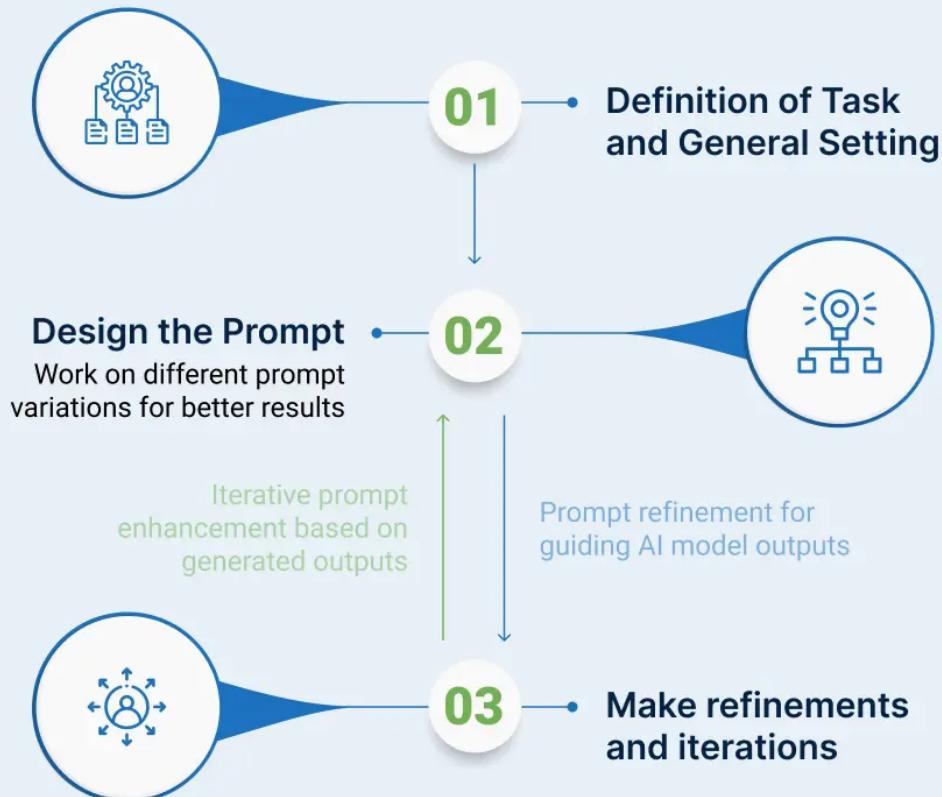


www.therundown.ai

What is prompt engineering?

Prompt Engineering is the process of designing, testing, and optimizing prompts that are used to instruct AI models to perform specific tasks.

Prompt Engineering Process



About prompt engineers

Prompt Engineer



The Role

- Work with cross-functional teams to discuss product development
- Identify uses of AI tools
- Design, develop and refine AI-generated text prompts

Background

- Bachelor's degree in Computer Science or Machine Learning or a related field
- Additional certifications recommended

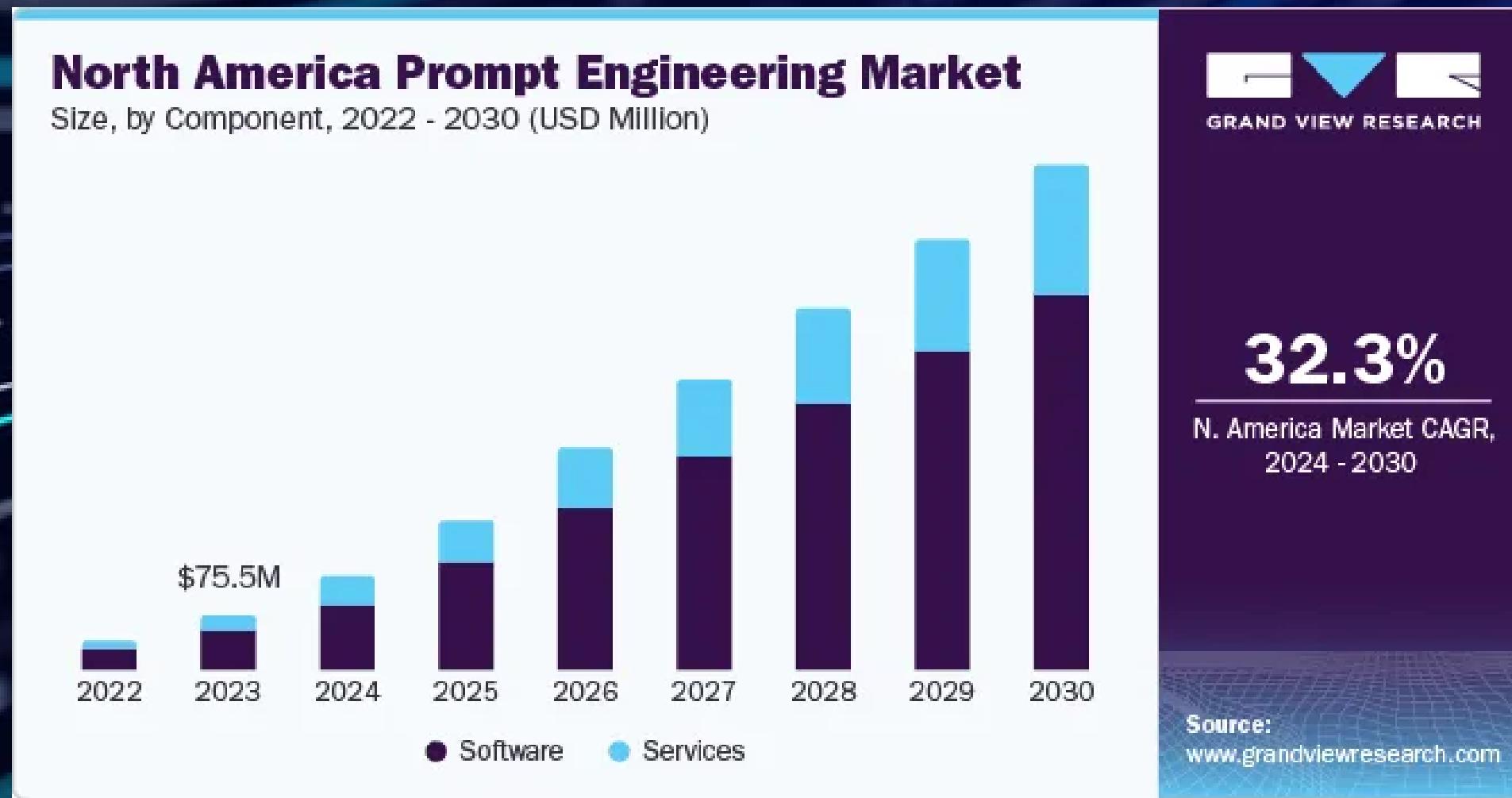
Skills

- Excellent knowledge of natural language processing
- Knowledge of machine learning
- Comprehensive knowledge of AI-generated content development

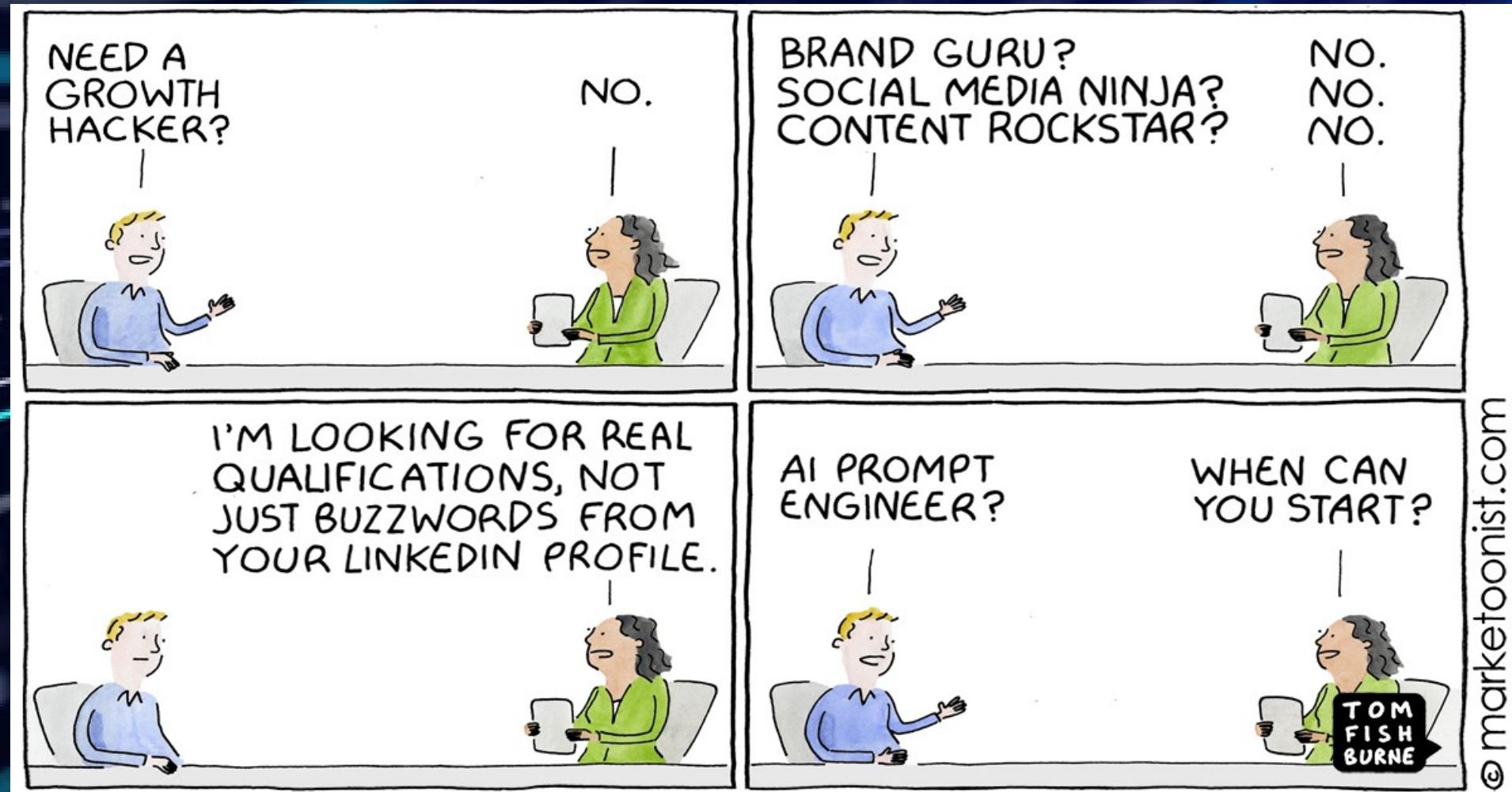
Salary

Junior: \$ 280,000
Average: \$ 327,000
Senior: \$ 375,000

Prompt engineers are highly valued.



Prompt engineers are in high demand.



Benefits and requirements of prompt engineers

<div class="caption"> Original post </div>

.com/AiBreakfast/status/1625594172815339521

← Post



AI Breakfast ✅
@AiBreakfast

Wow - Anthropic (Google's latest \$300M AI investment) is hiring "Prompt Engineer" for \$250k-\$335k/yr + equity

No CS degree required, just have "at least basic programming skills"

Wild times.

Representative Projects

- Discover, test, and document best practices for a wide range of tasks relevant to our customers.
- Build up a library of high quality prompts or prompt chains to accomplish a variety of tasks, with guide to help users search for the one that meets their needs.
- Build a set of tutorials and interactive tools that teach the art of prompt engineering to our customers.

You may be a good fit if you:

- Have a creative hacker spirit and love solving puzzles.

- Are an excellent communicator, and love teaching technical concepts and creating high quality documentation that helps out others.

Prompt
engineering
requires
soft skills.

5 non-tech prompt engineering skills



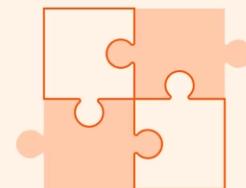
Language



Communication



Creativity



Critical thinking



Subject matter expertise

—zapier

Prompt engineering skill - language

Linguistics,

the scientific study of language and its structure, plays a vital role in prompt engineering. It helps in:

1. **Understanding Language Structure:** Helps in designing effective prompts.
2. **Semantic Analysis:** Understanding the meaning of words, phrases, and sentences helps in creating more precise prompts.
3. **Pragmatics:** Understanding the context in which language is used can lead to the creation of more effective prompts.

Prompt engineering skill - language

Prompt Engineering

How to structure Prompts:

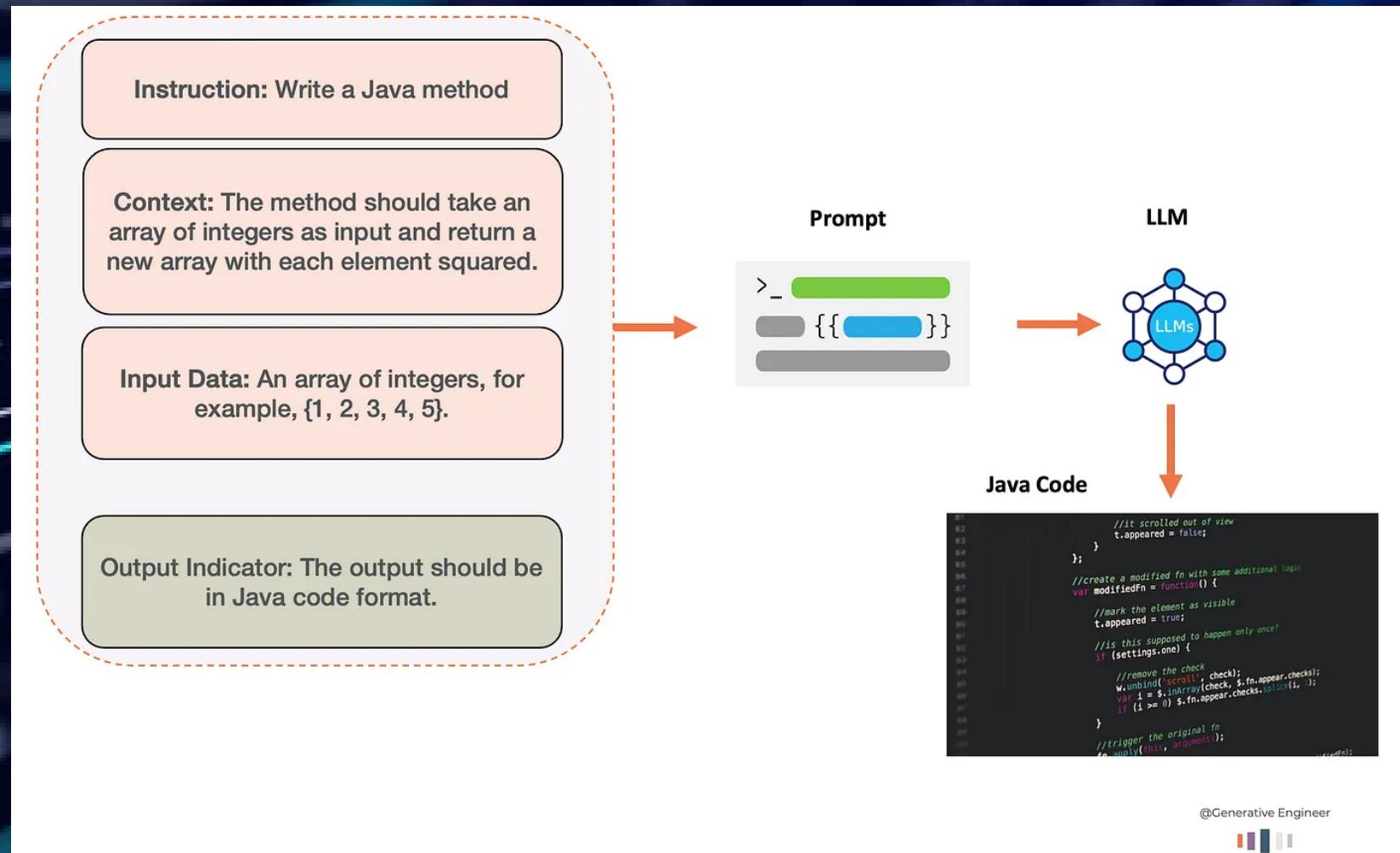
- Specify the scenario
- Use specific formulations
- Assign a role in the instruction
- Formulate the instructions clearly
- Provide the model with additional information

Prompt engineering skill - critical thinking

Prompting Approach	Description	Comparison to Human Thinking
Zero-shot Prompting	AI answers without prior examples, using its training	Like answering a question using only what you already know.
Few-shot Prompting	AI uses a few examples to understand how to respond.	Learning from a few examples before trying something new.
Chain-of-Thought Prompting	AI breaks down its process into steps before answering.	Thinking through steps to solve a problem, like in math.
Self-Consistency	AI generates multiple answers and picks the most consistent one.	Choosing the best solution after considering several options.
Generate Knowledge Prompting	AI creates new ideas or information.	Using imagination or knowledge to come up with new ideas.
Prompt Chaining	Sequential prompts guide AI through a process to a result.	Following steps in a plan to achieve a goal.
Active-Prompt	Prompts adjust based on AI's responses in real-time.	Adjusting conversation based on feedback, like in a dialogue.
Directional Stimulus Prompting	Prompts direct AI towards a desired outcome.	Organizing thoughts or information in a structured way.
Multimodal CoT	AI organizes information in graph structures for tasks.	Organizing thoughts or information in a structured way.

@Generative Engineer

Prompt engineering skill - coding



English as the most powerful programming language

santiagof.medium.com/english-is-the-most-powerful-programming-language-even-for-data-s... ☆

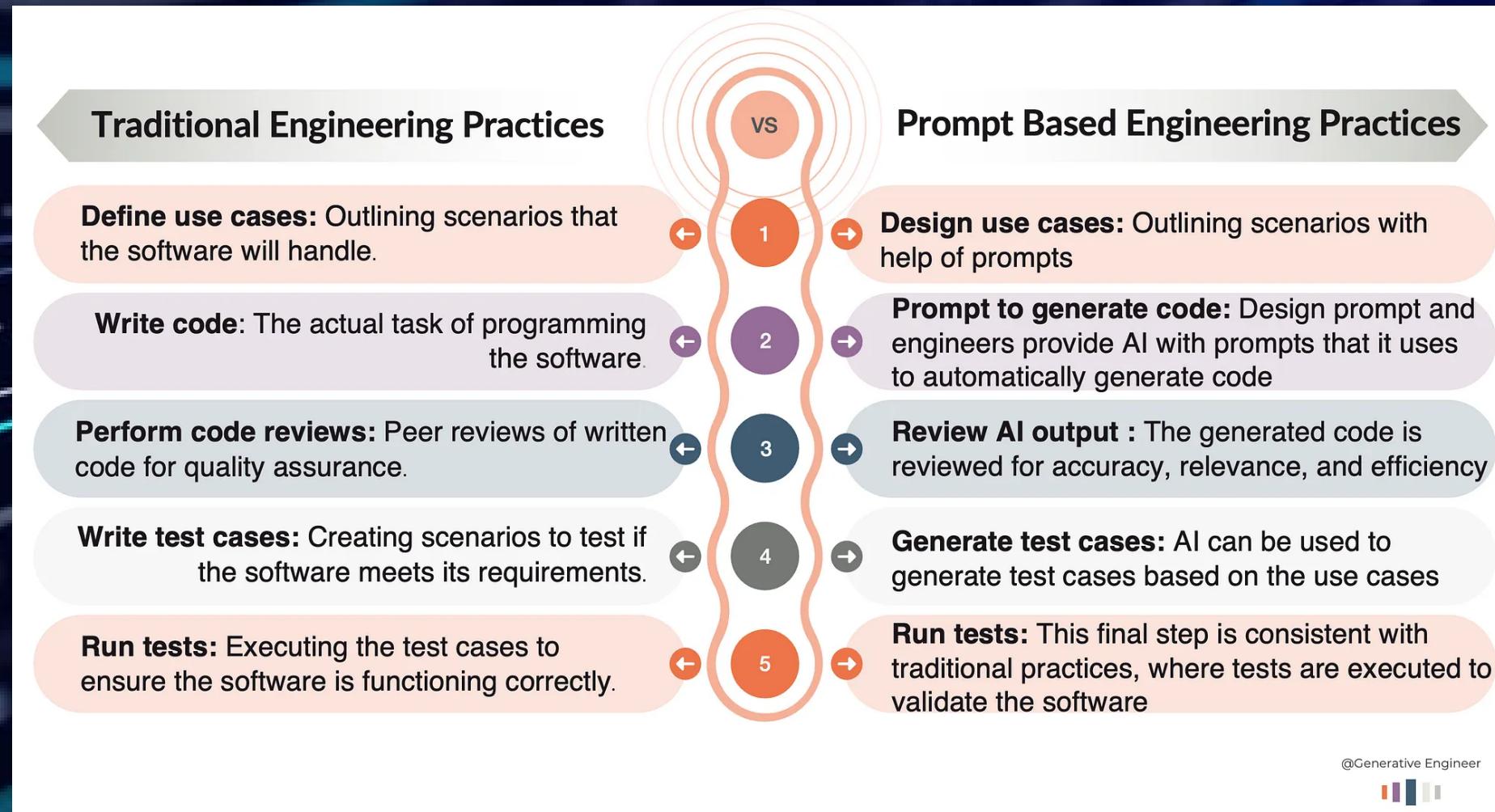
Member-only story

English is the most powerful programming language — even for data science: Introduction to prompt engineering

What prompt engineering is, which are the steps involved in it, and how it changes the way we solve problems with ML.

 Facundo Santiago · Follow
10 min read · May 9, 2023

GenAI creates engineering paradigm shift.

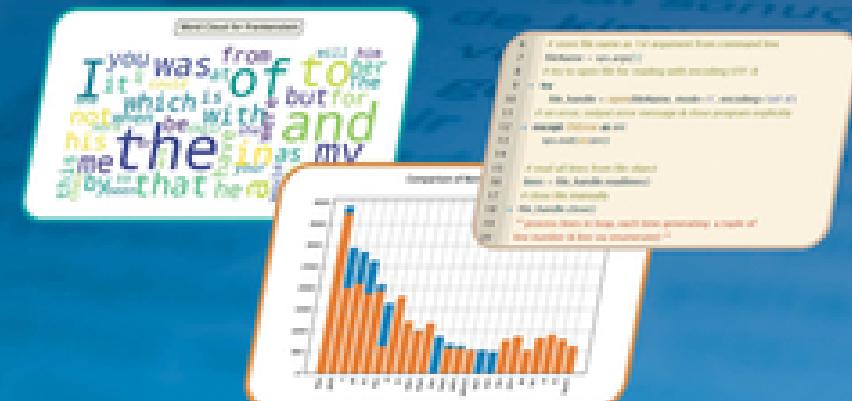


Start with Python  if
you wanna try
programming.

MARTIN WEISSE

Python Programming for Linguistics and Digital Humanities

APPLICATIONS FOR TEXT-FOCUSED FIELDS



WILEY Blackwell

Example: extracting ordered items from a text

Online order

Hey there! I'd like to make an order for pick-up. Could I get one large fries, two fish fillet sandwiches, three cheeseburgers with no onions, and four vanilla milkshakes? Oh, and could you also add five apple pies to that order? Thanks a lot!

Example: extracting ordered items from a text

Using knowledge of linguistics and common sense

```
[  
{'LIKE_NUM': True},  
{'POS': 'ADJ', 'OP': '?'},  
{'POS': 'NOUN', 'OP': '+'}]
```

Try it out [here!](#)

Outcome of the extraction

The screenshot shows the displaCy extraction interface with three stacked extraction trees. Each tree has a red minus sign at the top right and an 'add attribute' button at the bottom left.

- Top Tree:** Root node is `LIKE_NUM`. It has a green checkmark icon and an 'add attribute' button. Below it are two children: `POS` and `OP`.
- Middle Tree:** Root node is `ADJ`. It has a red minus sign at the top right and an 'add attribute' button at the bottom left. Below it are two children: `POS` and `?`.
- Bottom Tree:** Root node is `NOUN`. It has a red minus sign at the top right and an 'add attribute' button at the bottom left. Below it are two children: `POS` and `OP`.

At the bottom of the interface are two buttons: `add token` and `refresh text`.

Hey there! I'd like to make an order for pick-up. Could I get

one large fries , **two fish fillet sandwiches** ,
three cheeseburgers with no onions, and
four vanilla milkshakes ? Oh, and could you also add
five apple pies to that order? Thanks a lot!

Note that this demo currently doesn't indicate overlapping matches.

At the top right of this panel are three buttons: `Show tokens`, `displaCy` (with a question mark), and `displaCy ENT` (with a question mark).

Example: extracting ordered items from a text

Using a prompt

You work at a fast food restaurant and are good at summarizing what a customer orders from a text. Extract ordered items from the following text. Use the json format, with two keys, quantity and item:

Text:

<Hey there! I'd like to make an order for pick-up. Could I get one large fries, two fish fillet sandwiches, three cheeseburgers with no onions, and four vanilla milkshakes? Oh, and could you also add five apple pies to that order? Thanks a lot!>

Outcome of the extraction

```
[{"quantity": 1, "item": "large fries"},  
 {"quantity": 2, "item": "fish fillet sandwiches"},  
 {"quantity": 3, "item": "cheeseburgers with no onions"},  
 {"quantity": 4, "item": "vanilla milkshakes"},  
 {"quantity": 5, "item": "apple pies"}]
```

See the conversation on [Bing Chat!](#)

Take-away messages

3 points

- **Interdisciplinary integration:** Generative AI and prompt engineering are not just for tech enthusiasts or computer science majors. They can be incredibly useful tools for humanities majors as well.
- **Creativity amplified:** Generative AI can be seen as a tool to amplify human creativity, not replace it. Remember, the AI is just a tool, the real creativity comes from you.
- **Words as action:** The field of generative AI heavily relies on understanding and manipulating language. This is where knowledge of linguistics comes in. Understanding how language works can help you better utilize and even improve these tools.

Any questions ?

PLEASE DON'T

ASK QUESTIONS

Quickmeme.com



How to reach me

- <i class="fa-solid fa-envelope">
</i> Email
howard.haowen@gmail.com
- <i class="fa-solid fa-house">
</i> Webpage
<https://howard-haowen.rohan.tw>