AI, LLM & Prompt Engineering

- Speaker:

  Dr. Haowen Jiang
- Date:

  Oct 24th, 2024

Outline

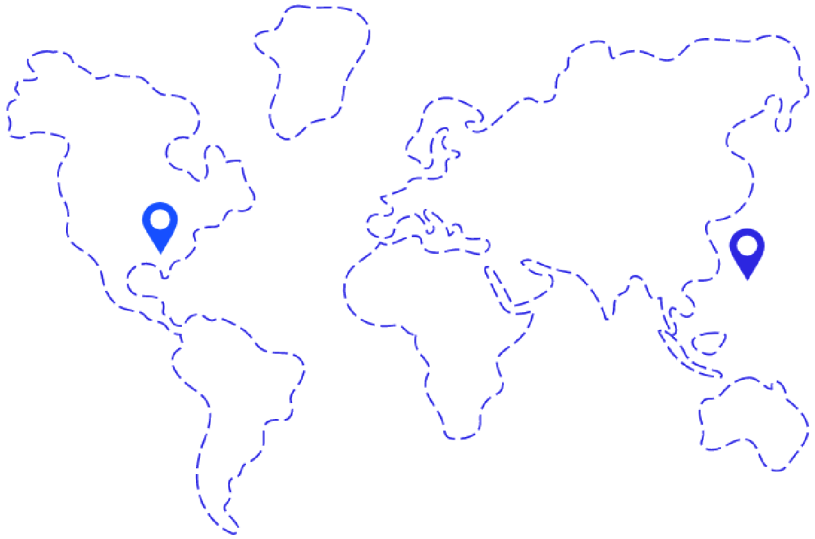
- My background
- Intro to AI
- Large Language Models (LLMs)
- Prompt Engineering
- Q & A


My background

Education

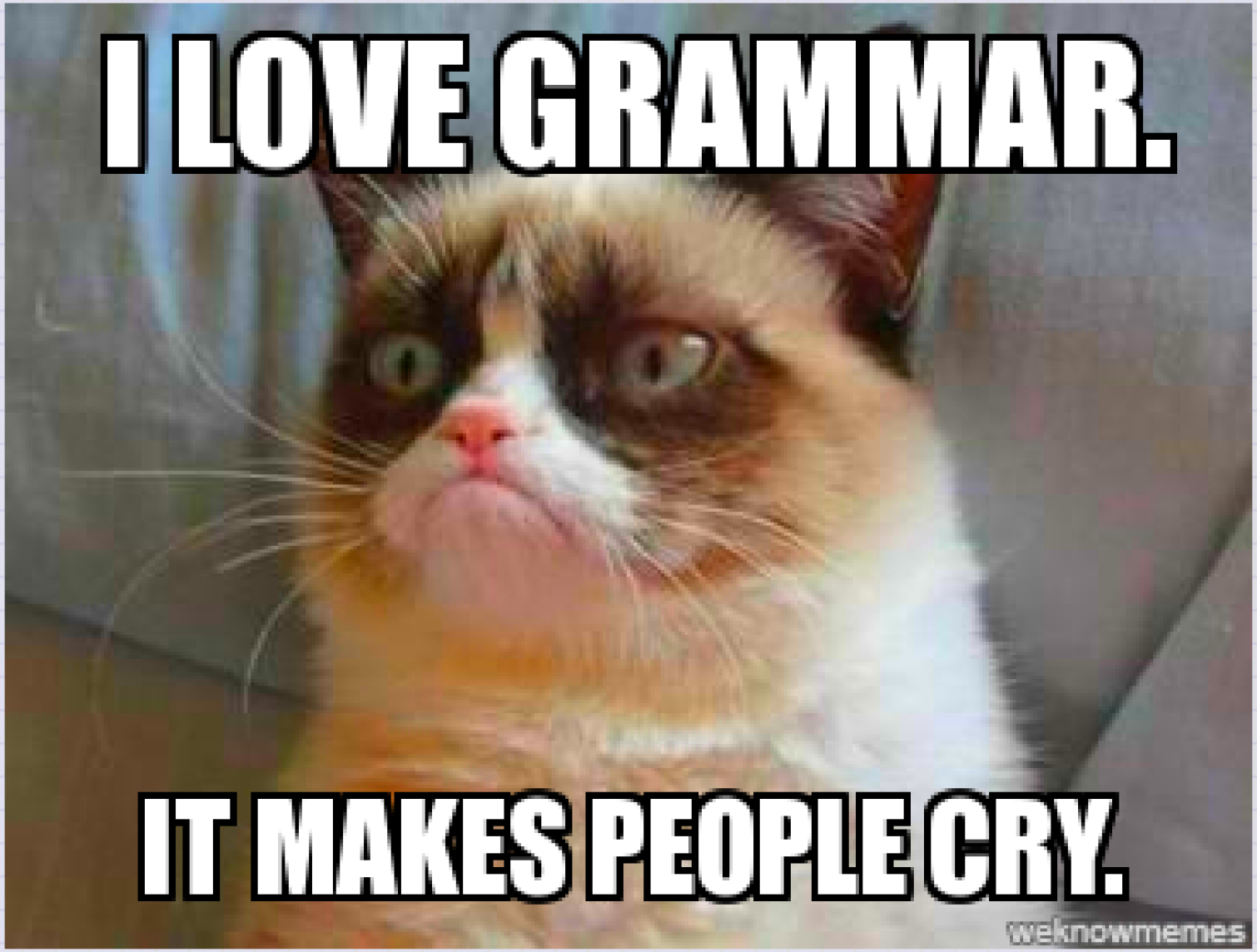
Nat'l Taiwan University
MA, Linguistics
BA, Foreign Languages

Rice University
PhD, Linguistics
MA, Linguistics



Experience

- 👩‍🏫 English lecturer at *Nat'l Taipei University of Technology*

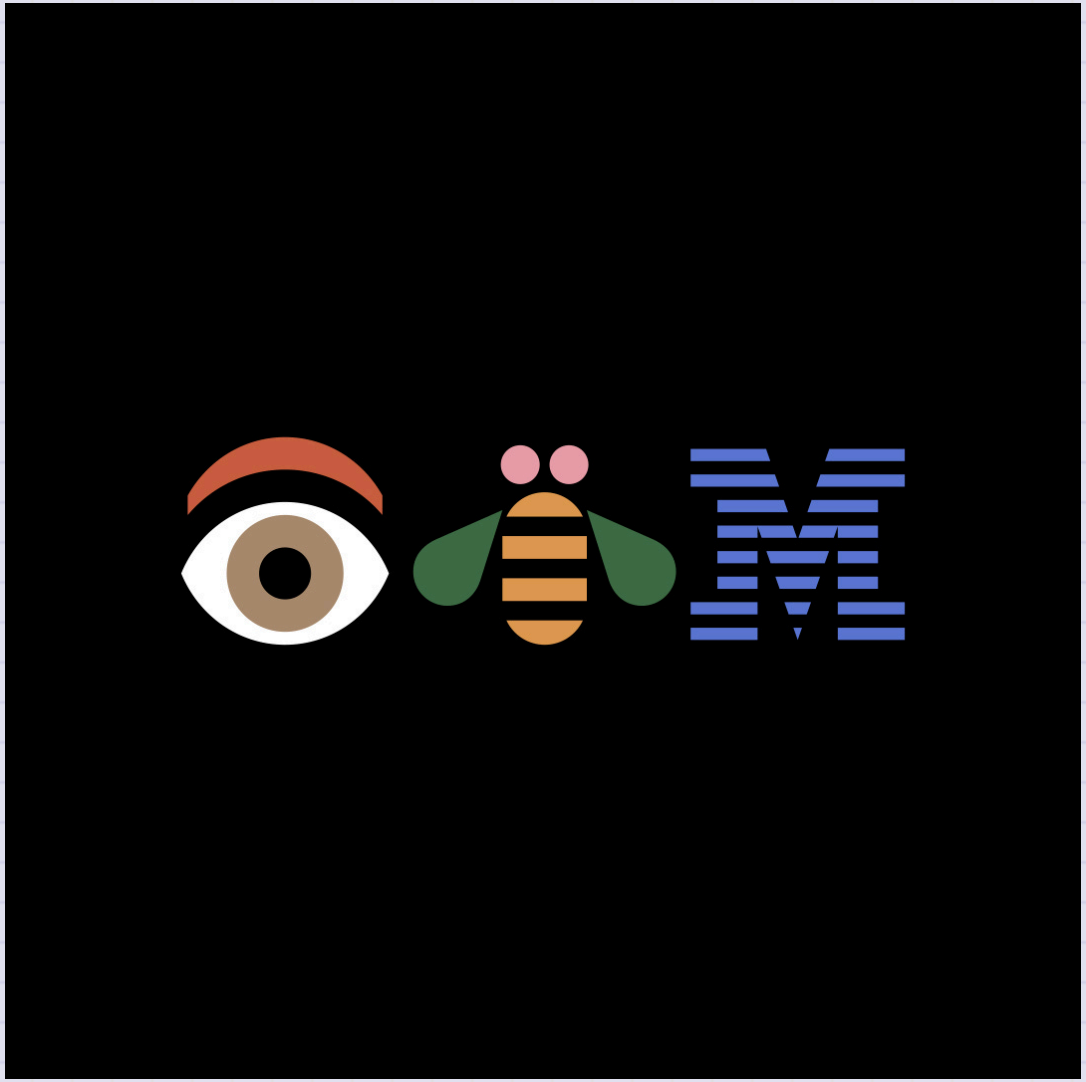


Experience

- 👩‍🔬 Postdoc researcher at *Peking University*

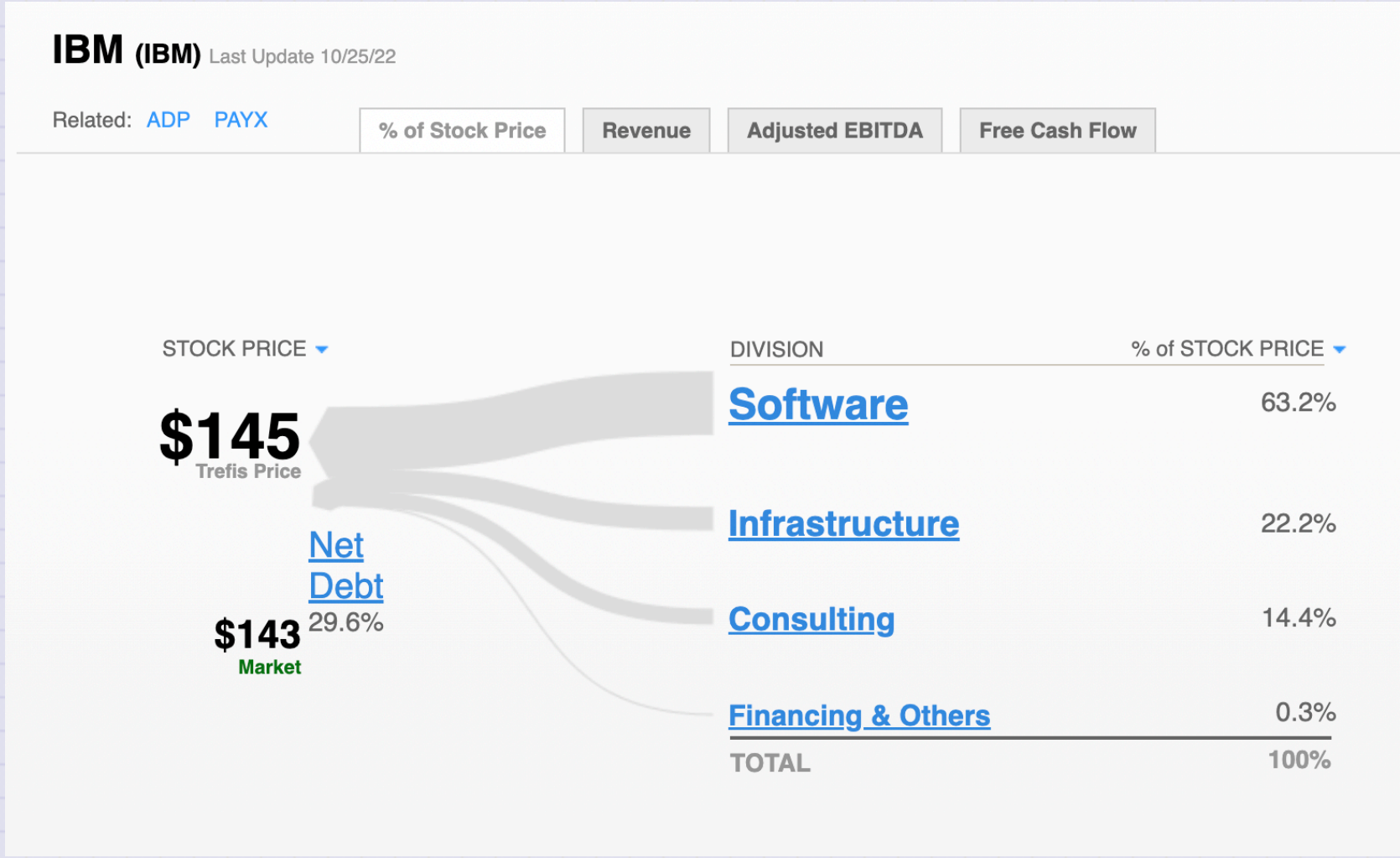
Experience

- 👨‍💻 AI engineer at IBM (my 5th industry job)



IBM Business Divisions

>> source




Intro to AI

Q1: What is AI? Or what counts as AI?

Q2: Do you think AI thinks as humans do?

DIFFERENCES BETWEEN
ARTIFICIAL INTELLIGENCE
MACHINE LEARNING & DEEP LEARNING

Artificial Intelligence

A broad concept that involves creating machines that can think and act like humans

Machine Learning

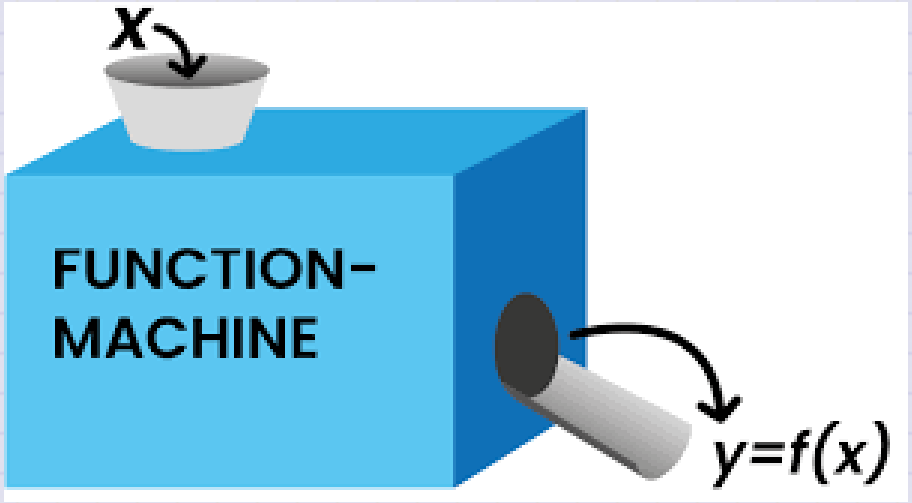
A subset of AI that focuses on creating algorithms that enable computers to learn from data and improve their performance over time.

Deep Learning

A subset of machine learning that focuses on neural networks with many layers.

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

- Natural Language

  `We like languages` >>> LIKE(We, languages)

- Programming

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

like("We", "languages")
```

Try it out >> here!

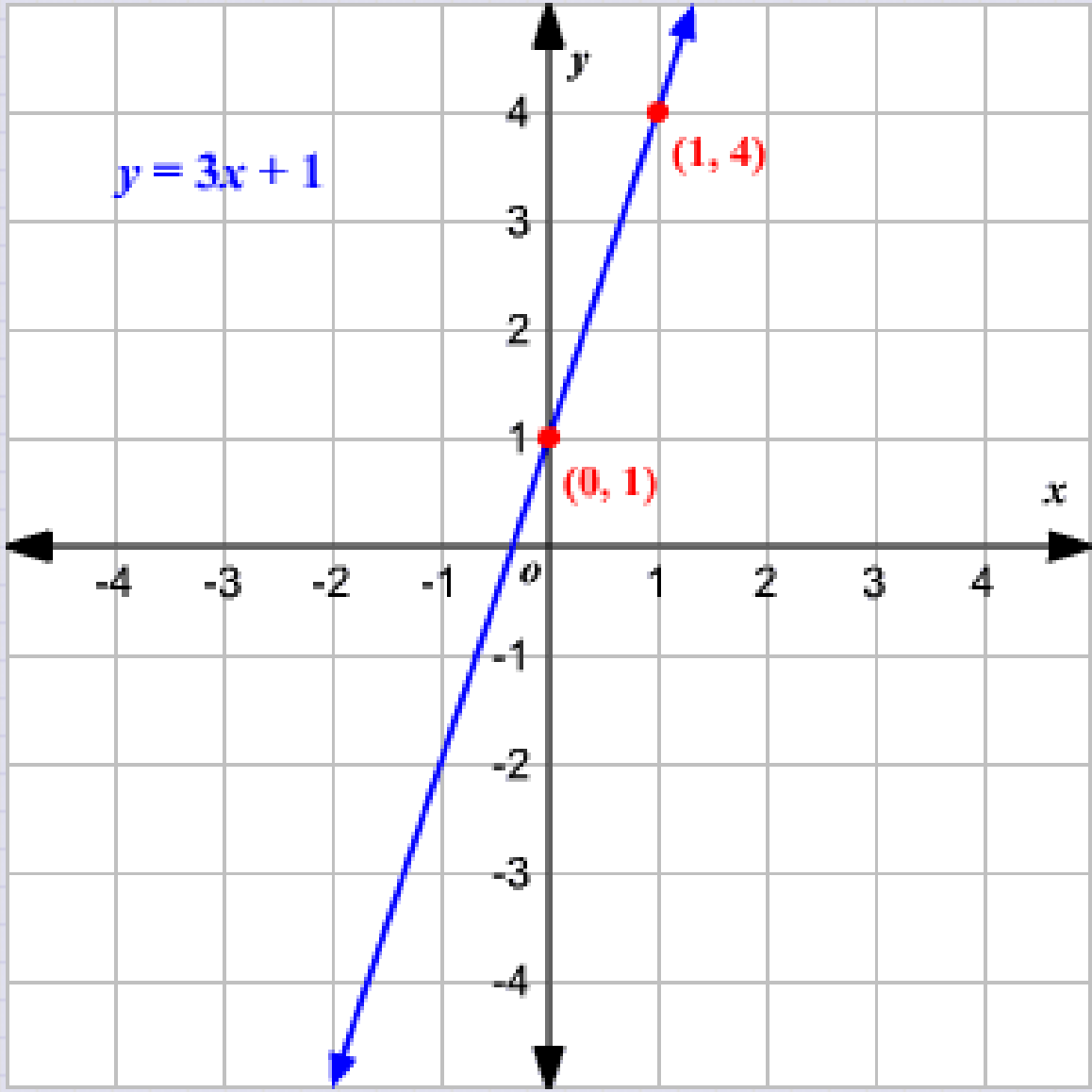
Data-driven AI

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

```
inputs = [0, 1, 2] # the x variable
outputs = [1, 4, 7] # the y variable
```

Random: y = 0.1x + 0.5

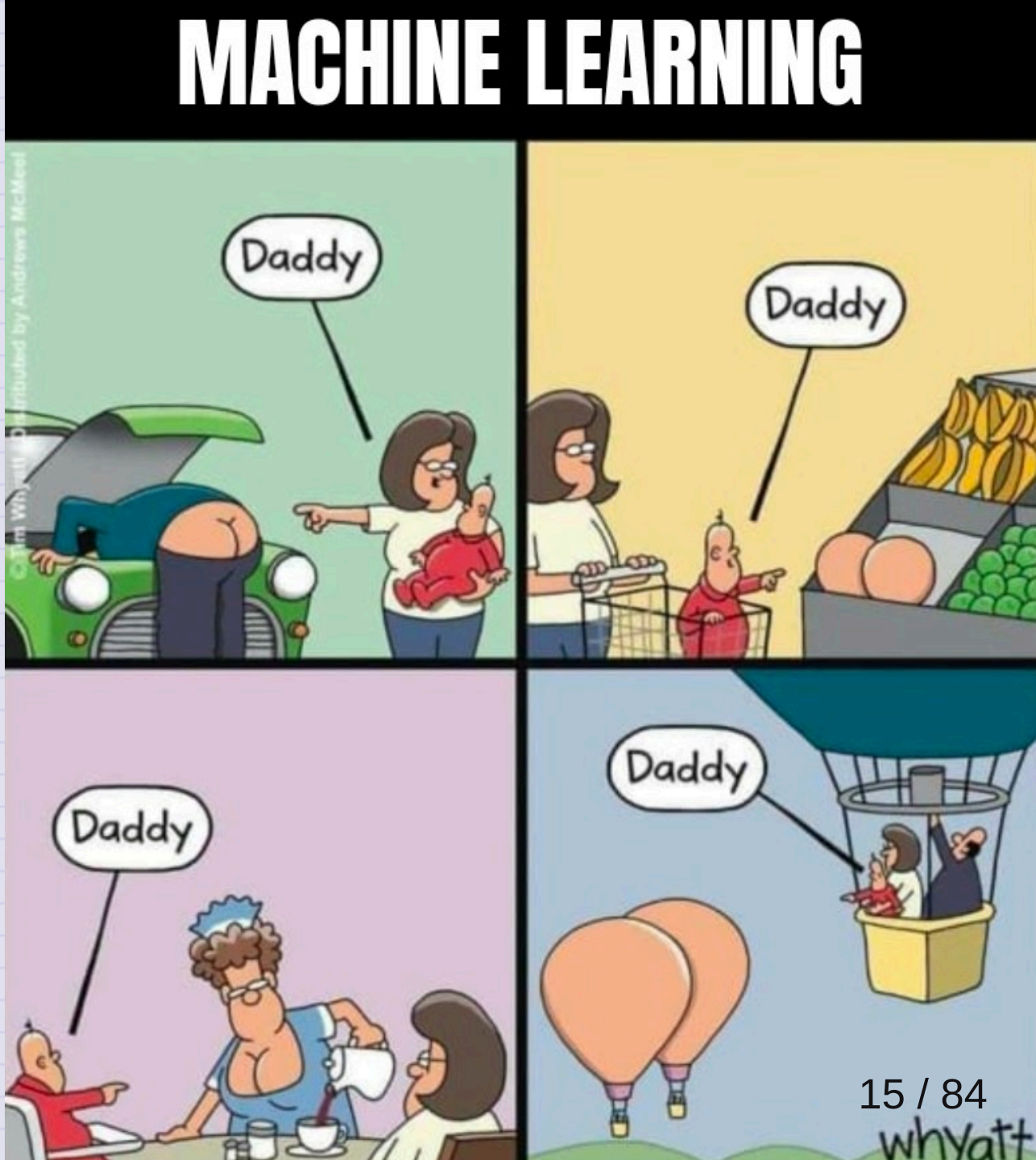
Trained: y = 3x + 1



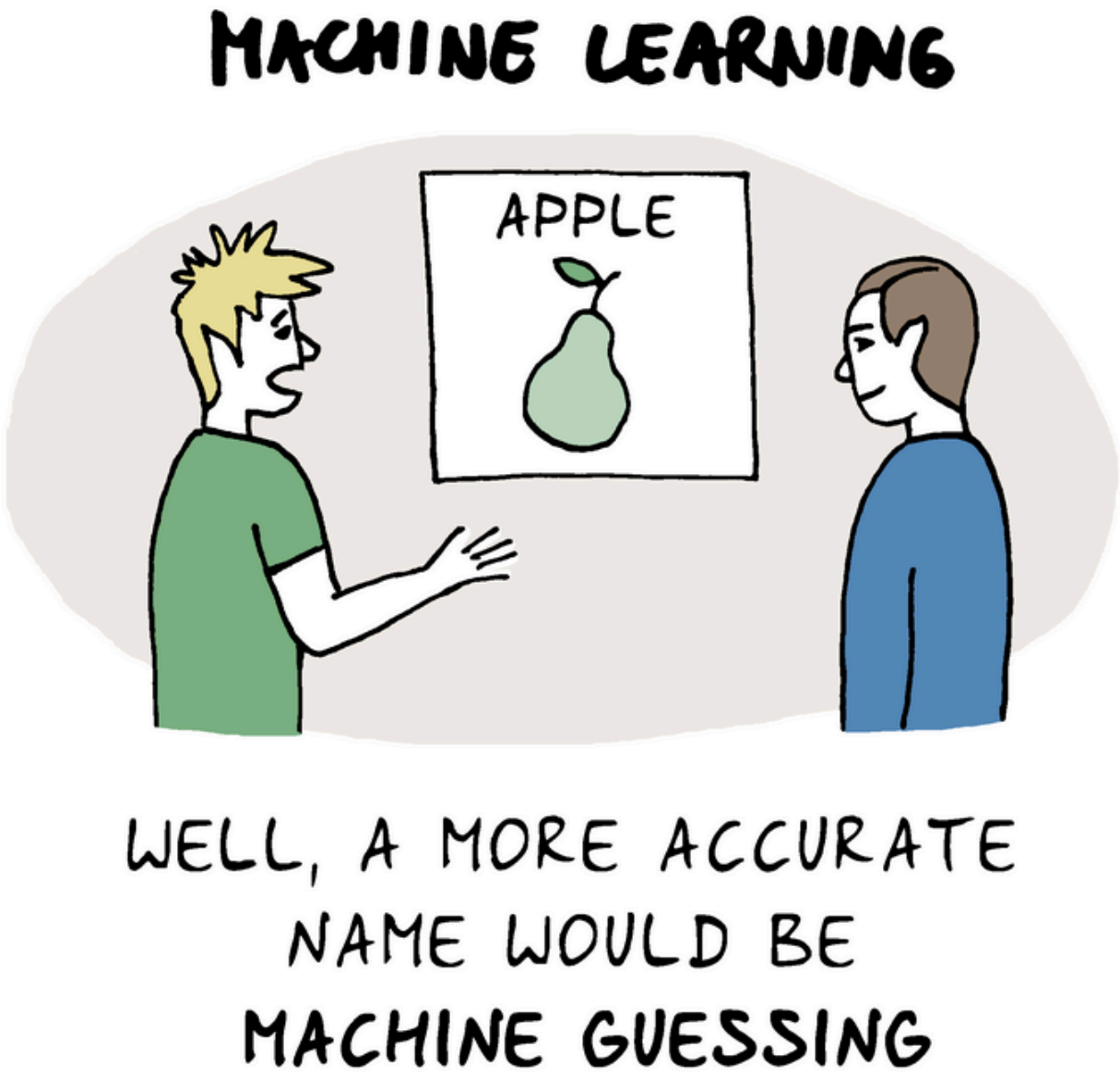
Haowen Jiang

How do machines learn?

> Machines learn by **trial and error**, just as humans do.



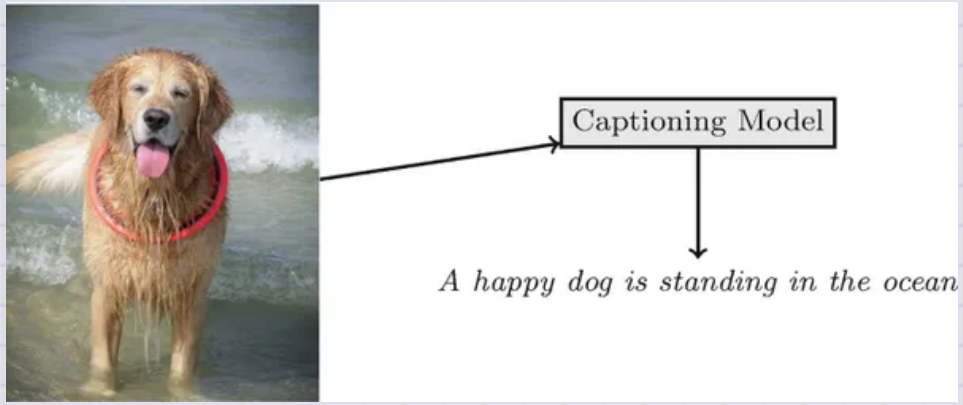
Machine Guessing



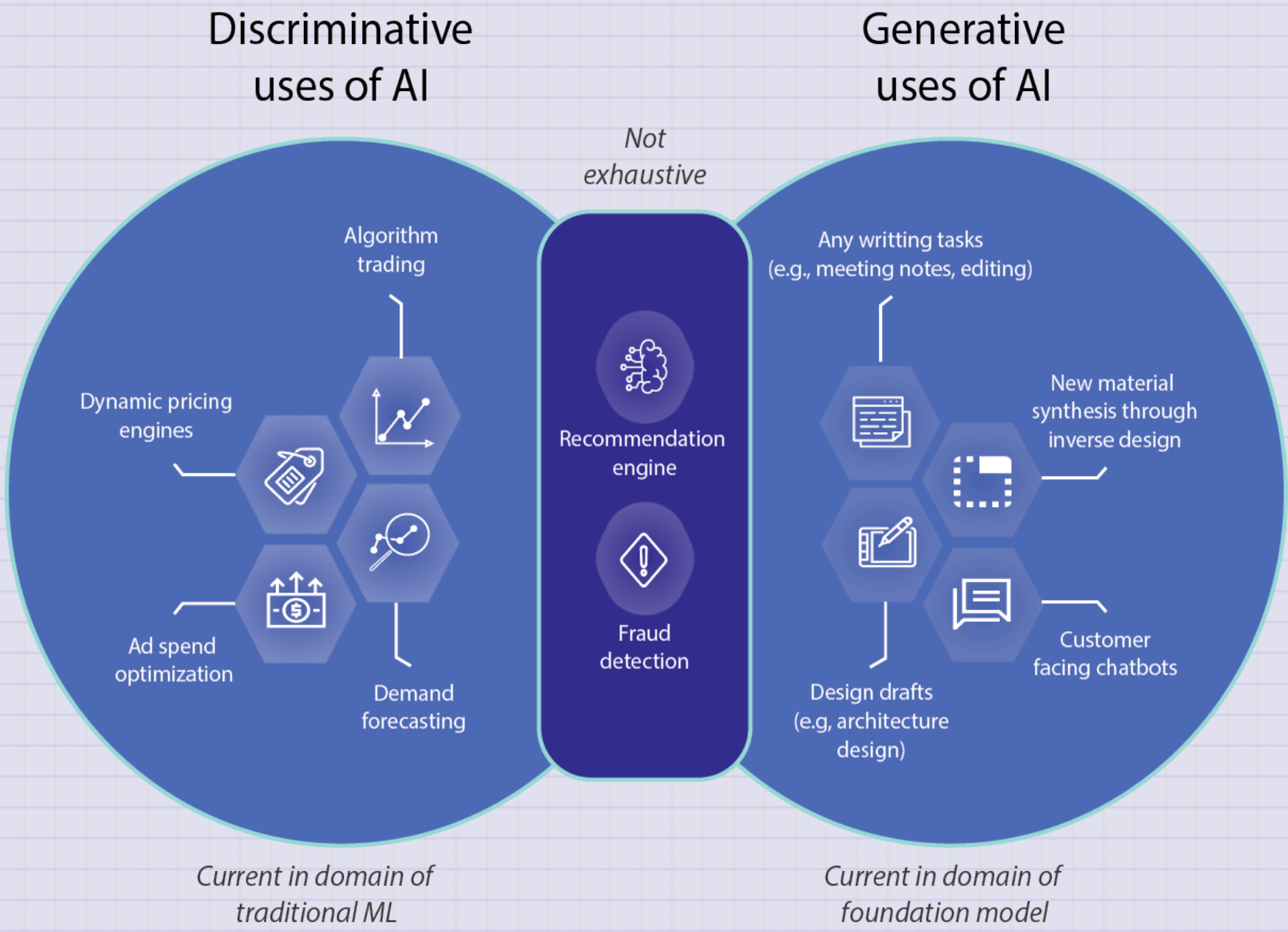
Why is machine learning powerful?

The true power of ML/DL lies in the fact that

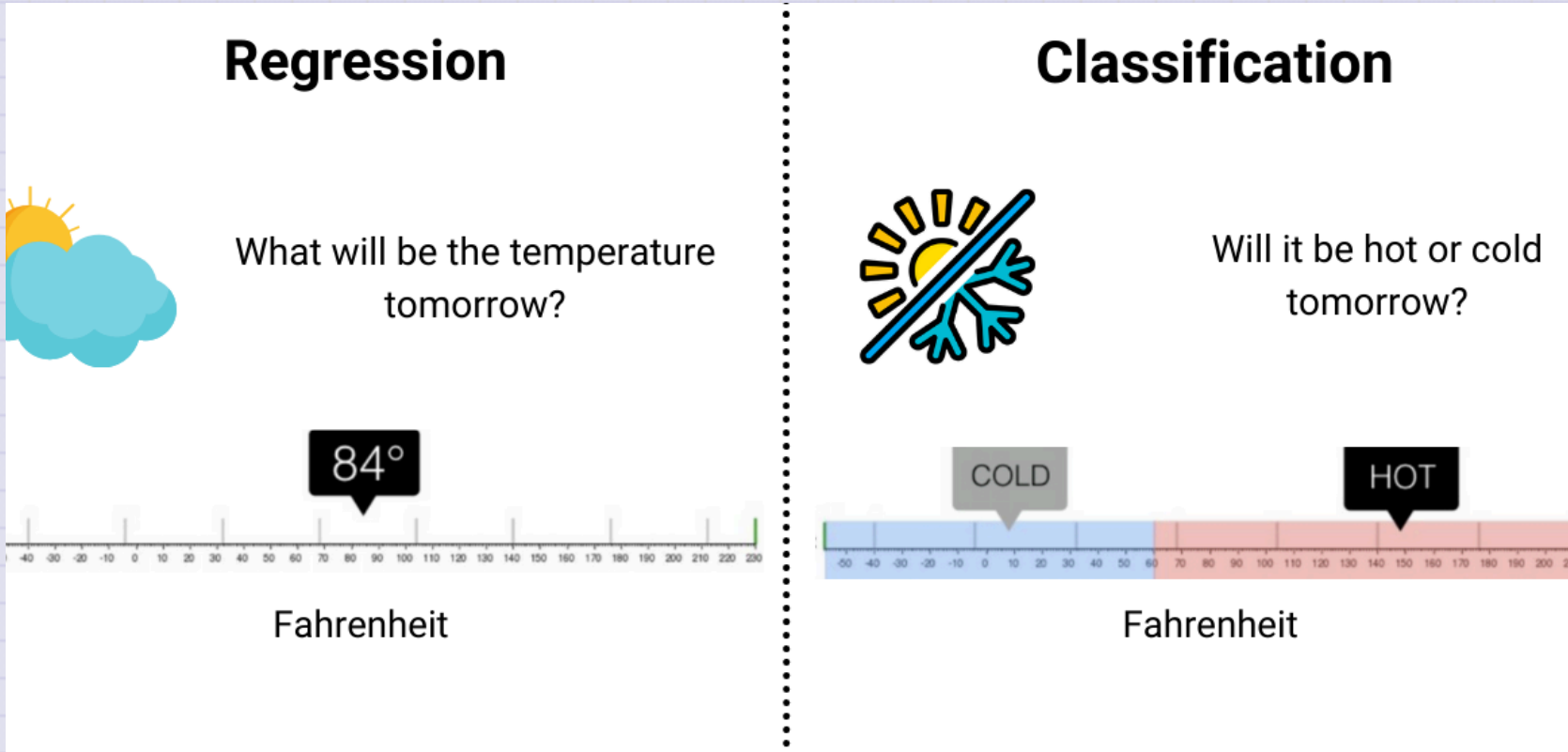
- humans only need to provide input and output
- computers are in charge of figuring out the right process, which is called a **model**



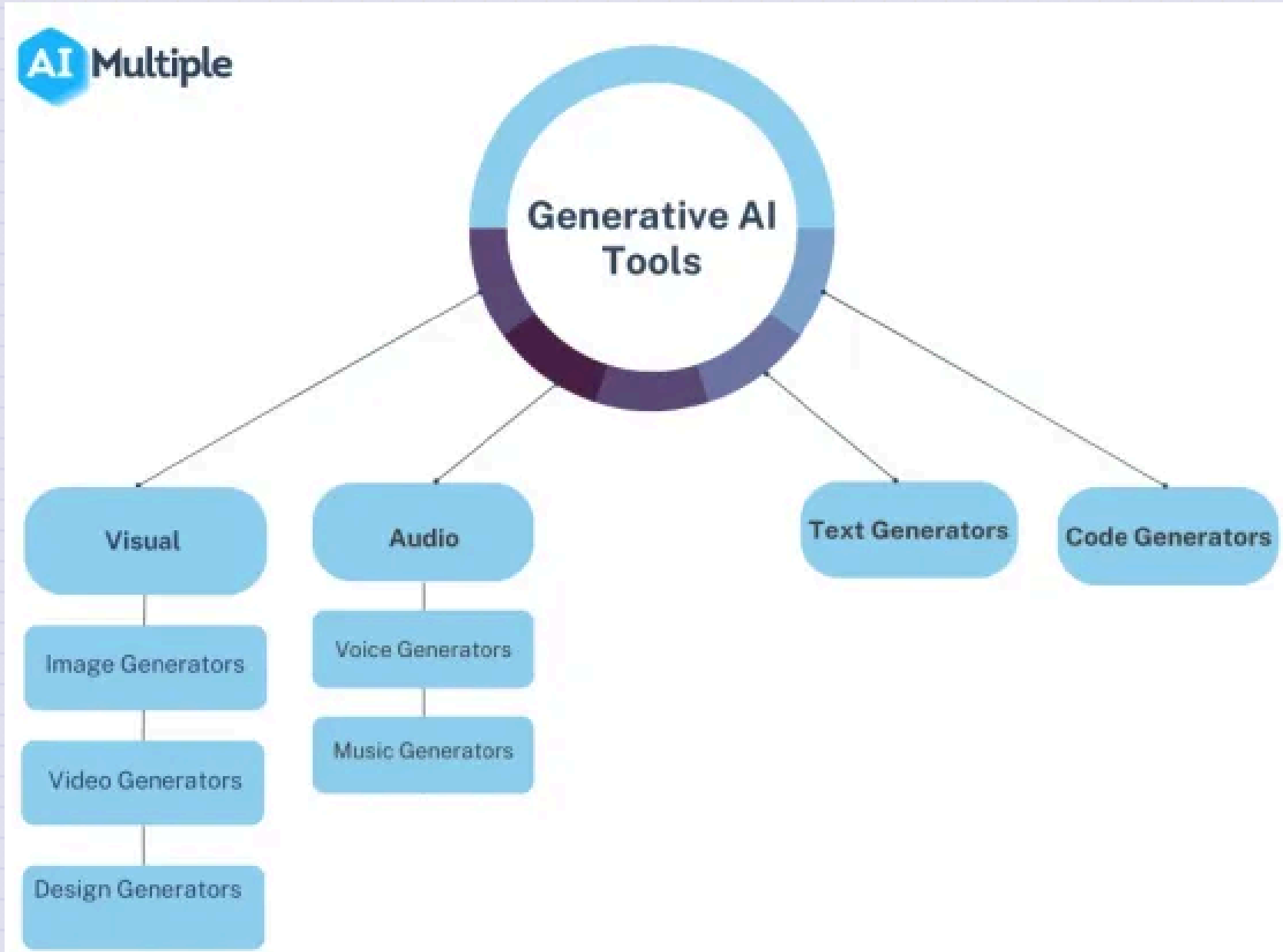
- 2 types of AI


Discriminative uses of AI
Generative uses of AI
Not exhaustive
Algorithm trading
Dynamic pricing engines
Ad spend optimization
Demand forecasting
Recommendation engine
Fraud detection
Any writting tasks (e.g., meeting notes, editing)
New material synthesis through inverse design
Customer facing chatbots
Design drafts (e.g, architecture design)
Current in domain of traditional ML
Current in domain of foundation model

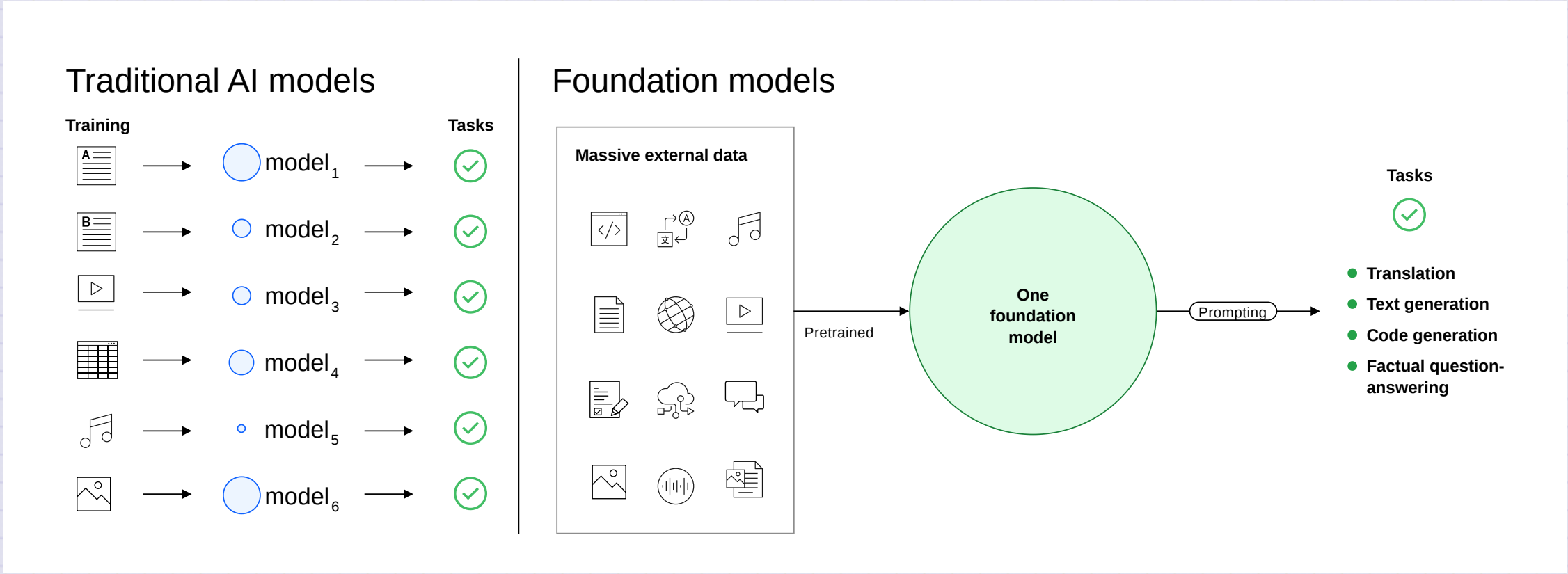
- 2 types of discriminative models



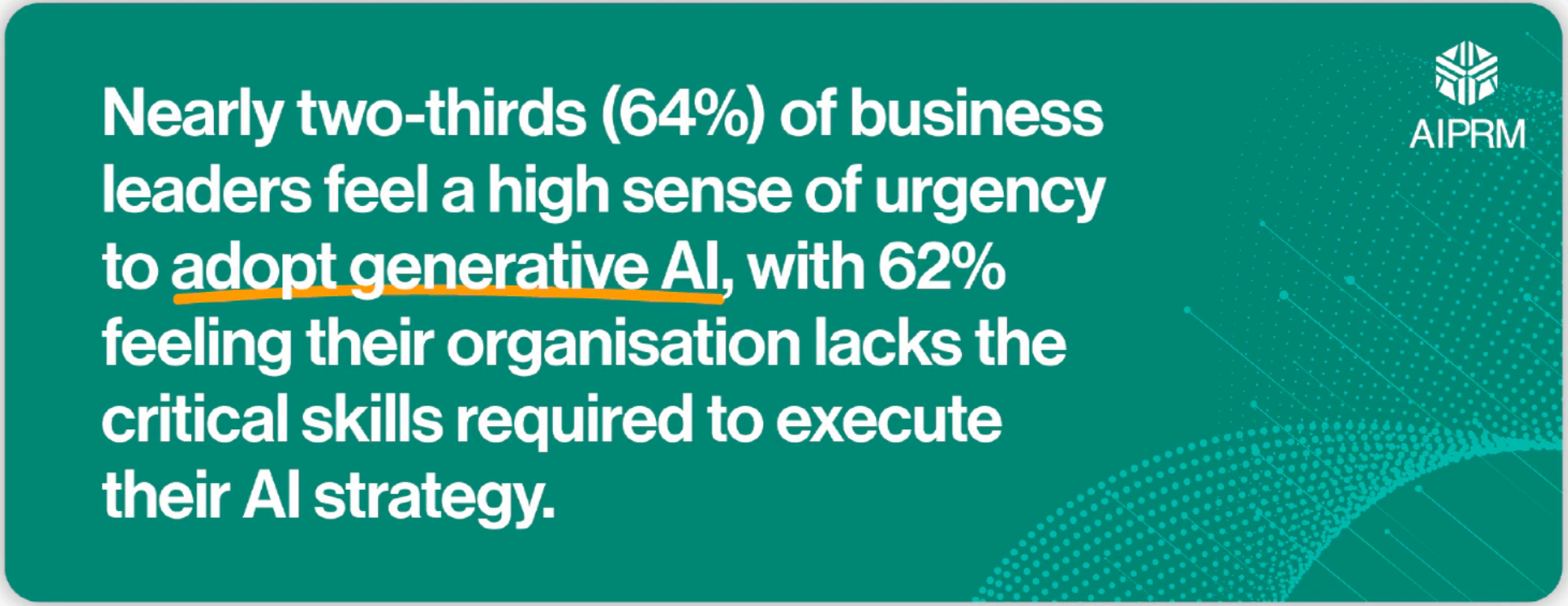
- 4 types of generative models



- Traditional AI vs Generative AI

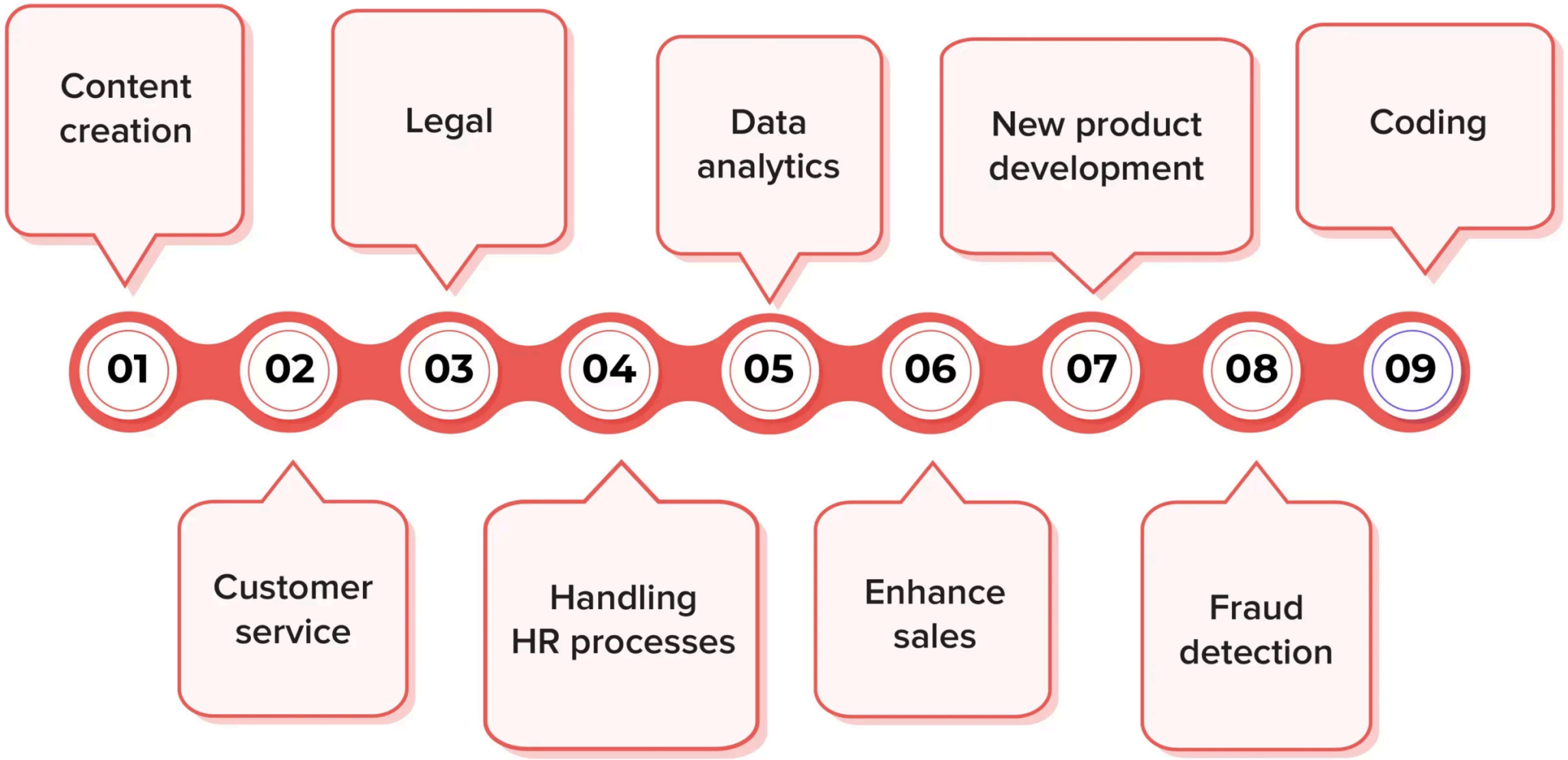


>> source



Source: AIPRM

Use cases of generative AI for business

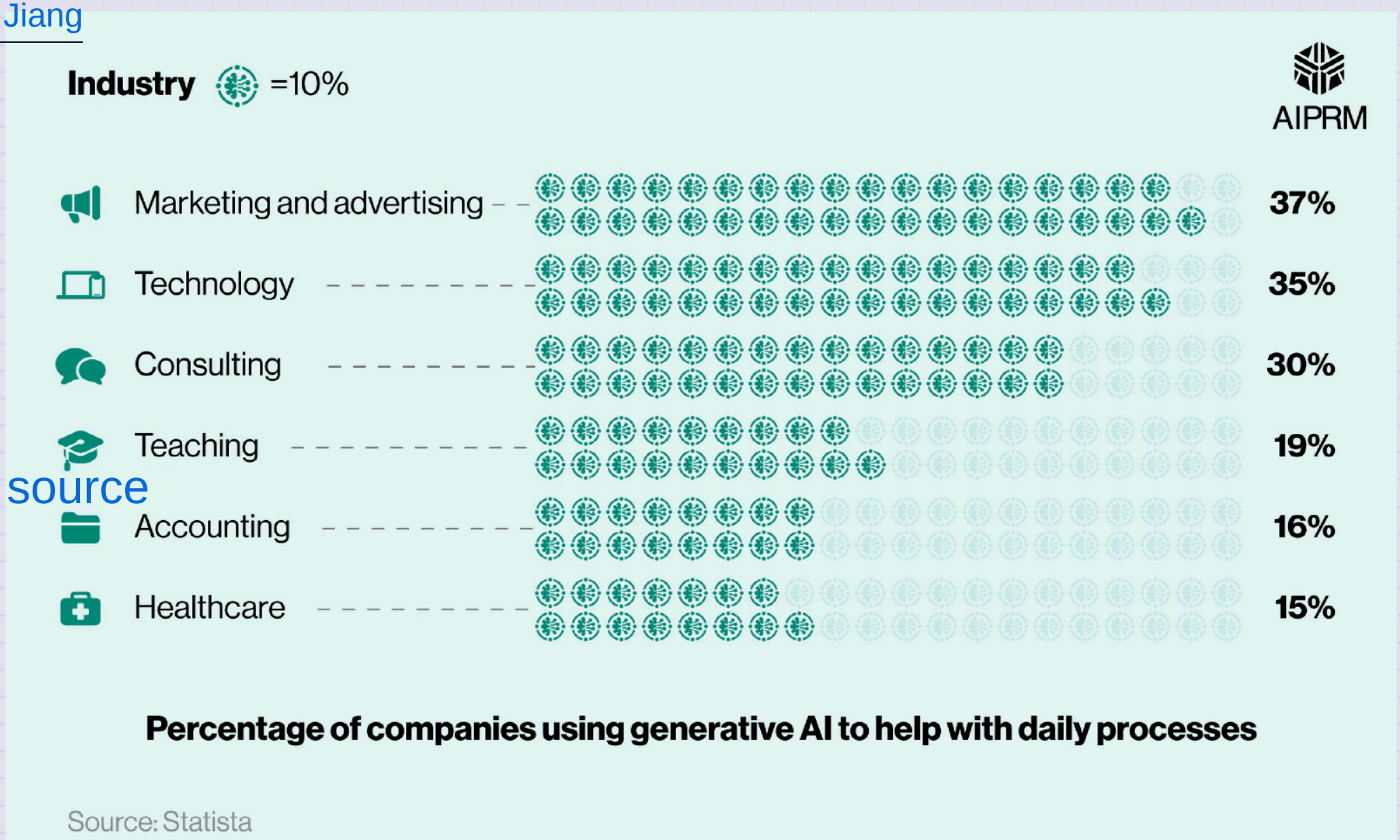


Applications of Generative AI in the Education Industry

>> source

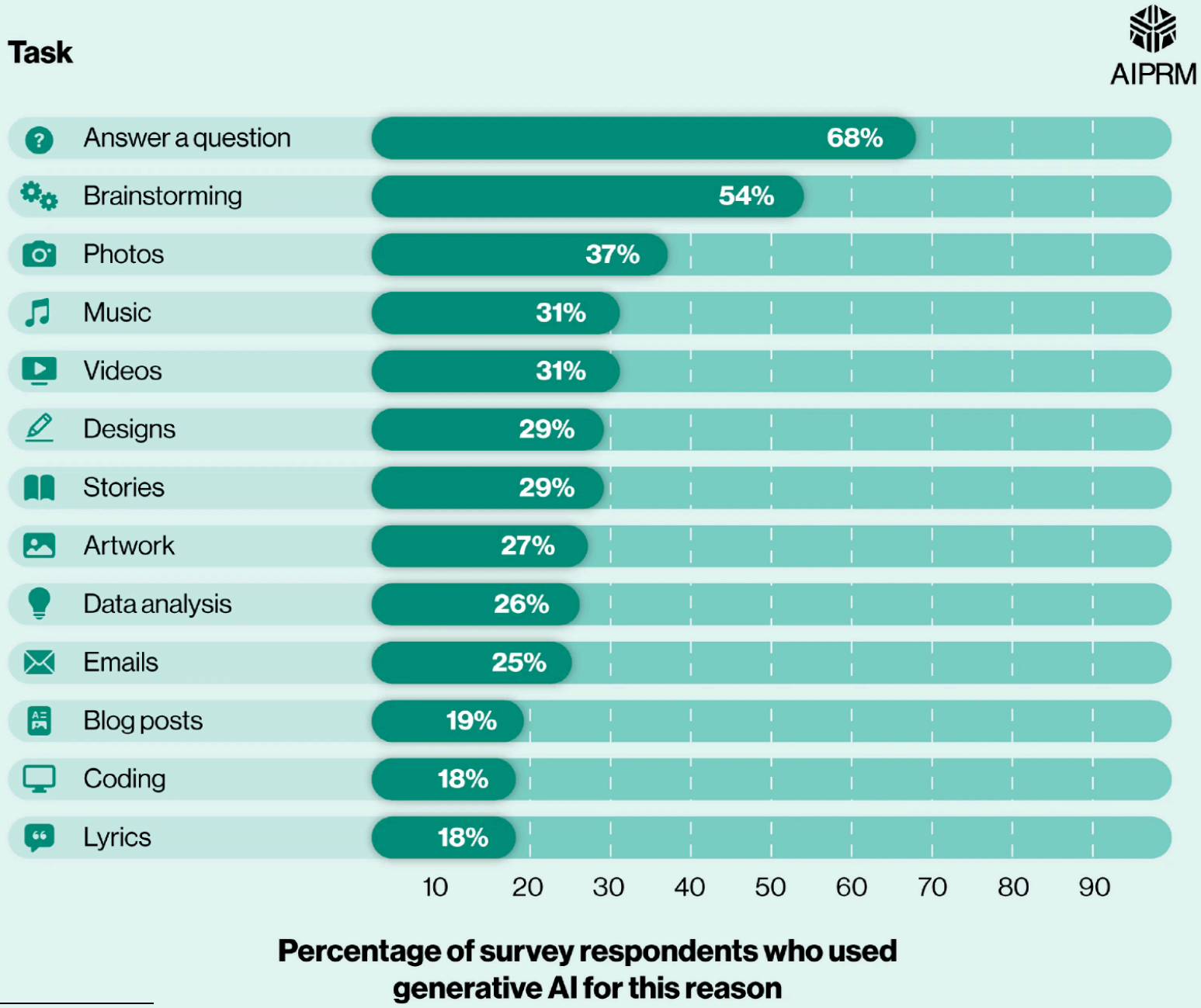
- Adaptive Learning
- Content Generation
- Automated Grading
- Personalized Tutoring
- Virtual Simulations
- Intelligent Learning System

XENONSTACK

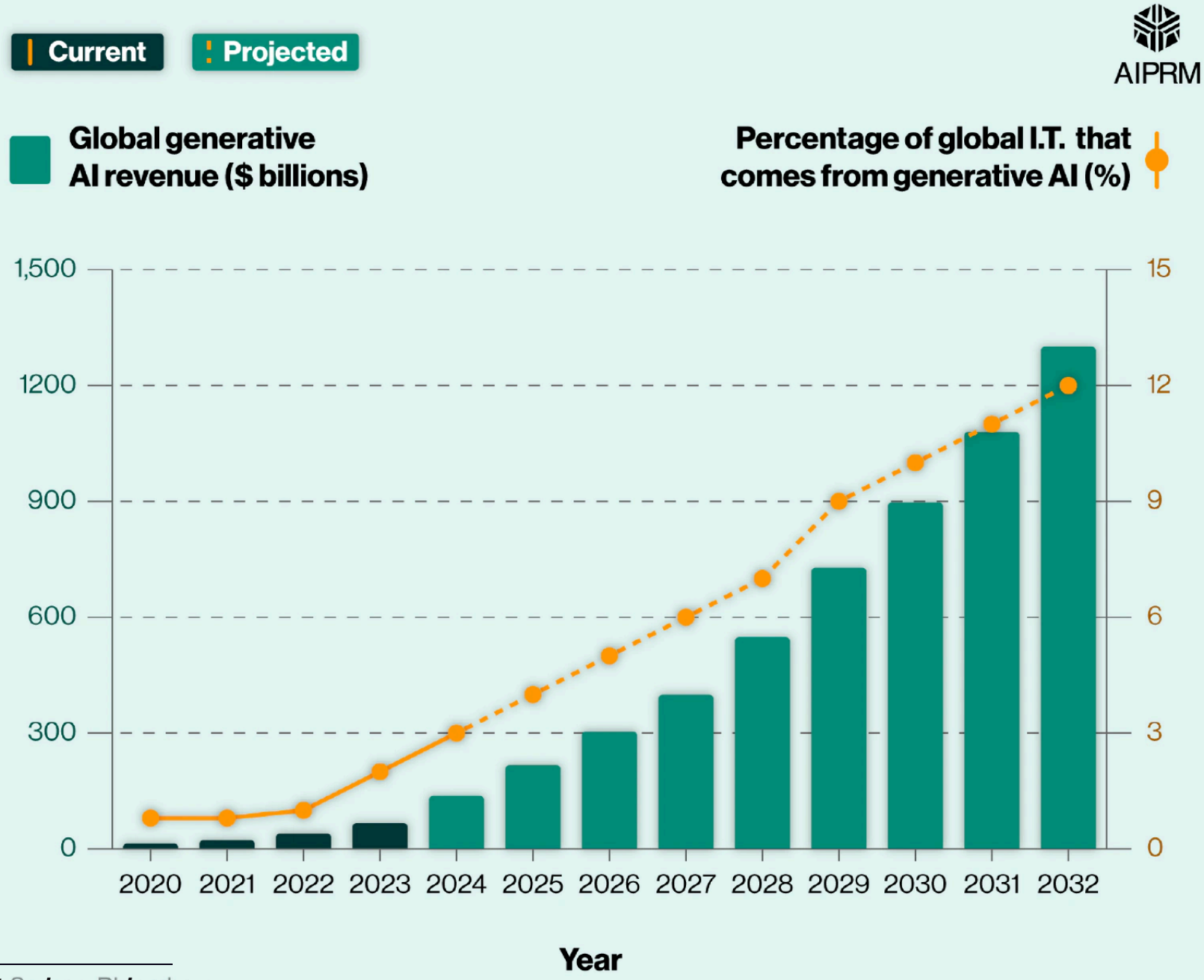

Industry =10%
AIPRM
Marketing and advertising
37%
Technology
35%
Consulting
30%
Teaching
19%
Accounting
16%
Healthcare
15%
Percentage of companies using generative AI to help with daily processes
Source: Statista

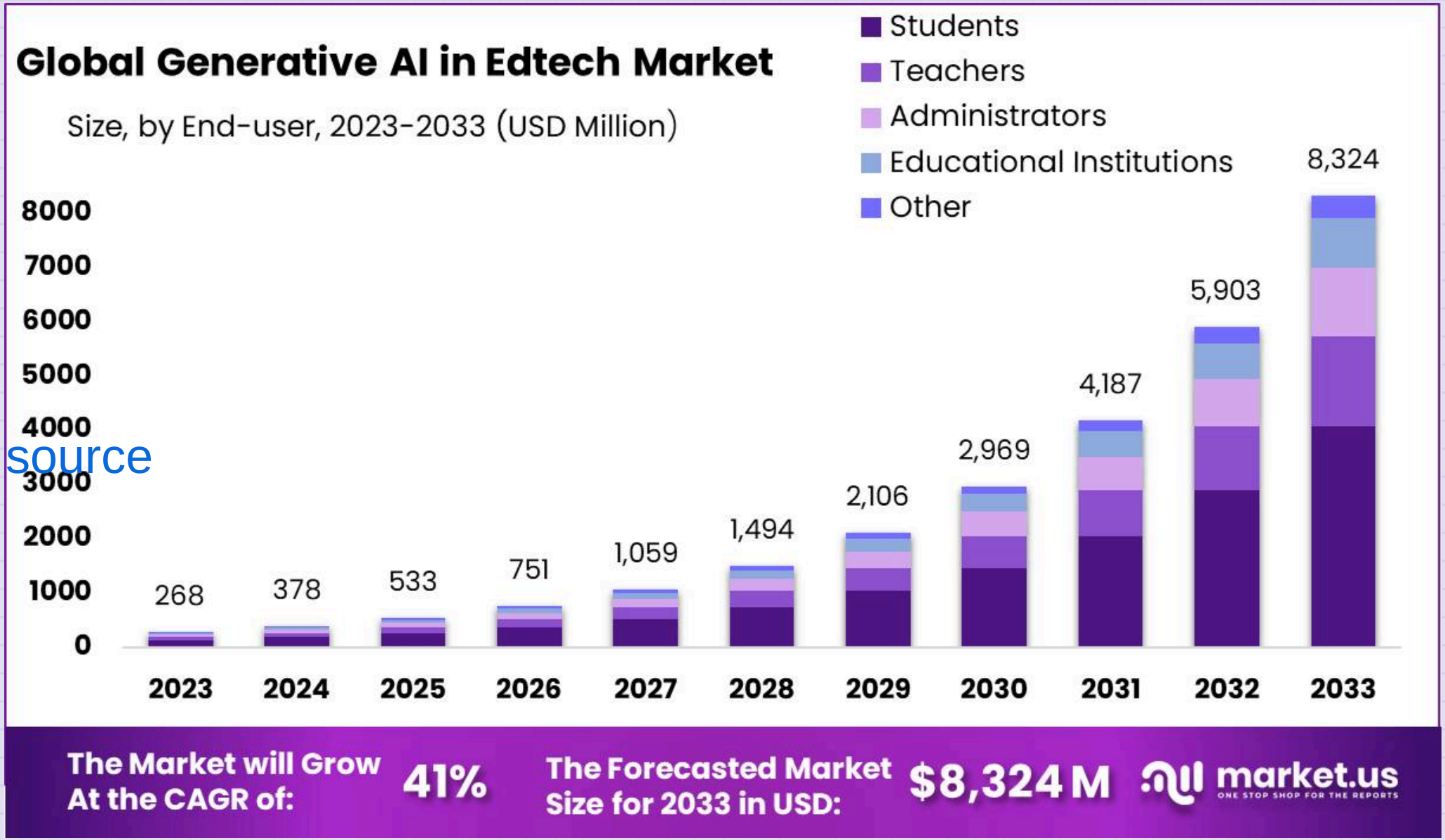
>> source

>> source


Task
AIPRM
Answer a question 68%
Brainstorming 54%
Photos 37%
Music 31%
Videos 31%
Designs 29%
Stories 29%
Artwork 27%
Data analysis 26%
Emails 25%
Blog posts 19%
Coding 18%
Lyrics 18%
10 20 30 40 50 60 70 80 90
Percentage of survey respondents who used generative AI for this reason

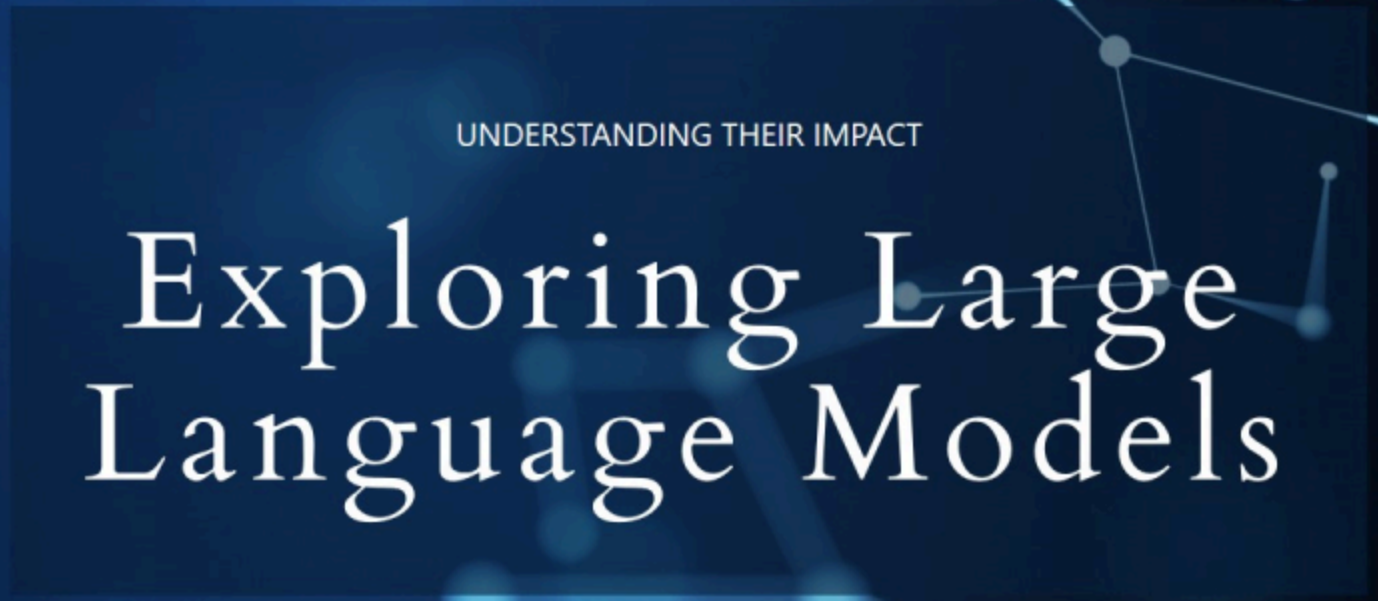
>> source


Current
Projected
AIPRM
Global generative AI revenue ($ billions)
Percentage of global I.T. that comes from generative AI (%)
1,500
1200
900
600
300
0
15
12
9
6
3
0
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
Year
Source: Bloomberg


Global Generative AI in Edtech Market
Size, by End-user, 2023-2033 (USD Million)
Students
Teachers
Administrators
Educational Institutions
Other
8000
7000
6000
5000
4000
3000
2000
1000
0
268
378
533
751
1,059
1,494
2,106
2,969
4,187
5,903
8,324
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
The Market will Grow At the CAGR of: 41%
The Forecasted Market Size for 2033 in USD: $8,324 M
market.us
ONE STOP SHOP FOR THE REPORTS

>> source

Large Language Models (LLMs)



Q1: What LLM APPs (ChatGPT, Bing Chat, Google Gemini) have you ever used?

Q2: What're some of the challenges you face when interacting with them?

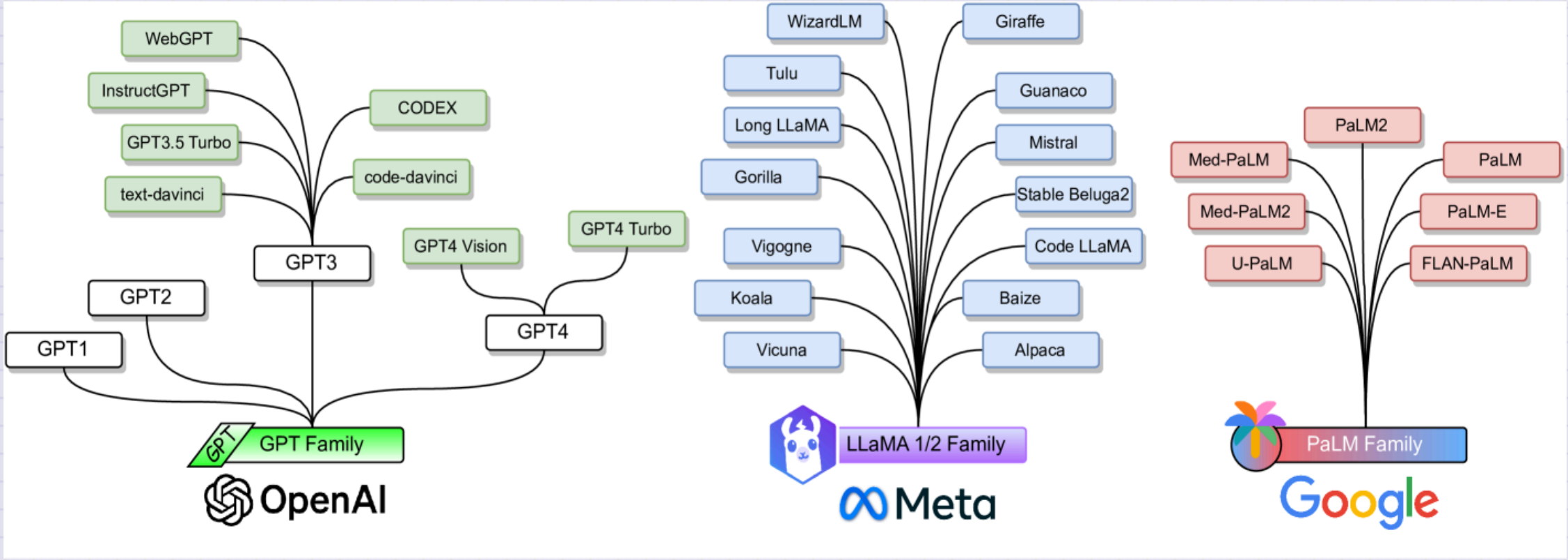
ChatGPT's significance

> >> TechGoing
>
> ChatGPT’s history is as significant as the birth of **the PC or the Internet**.
> ~ Bill Gates

ChatGPT's model

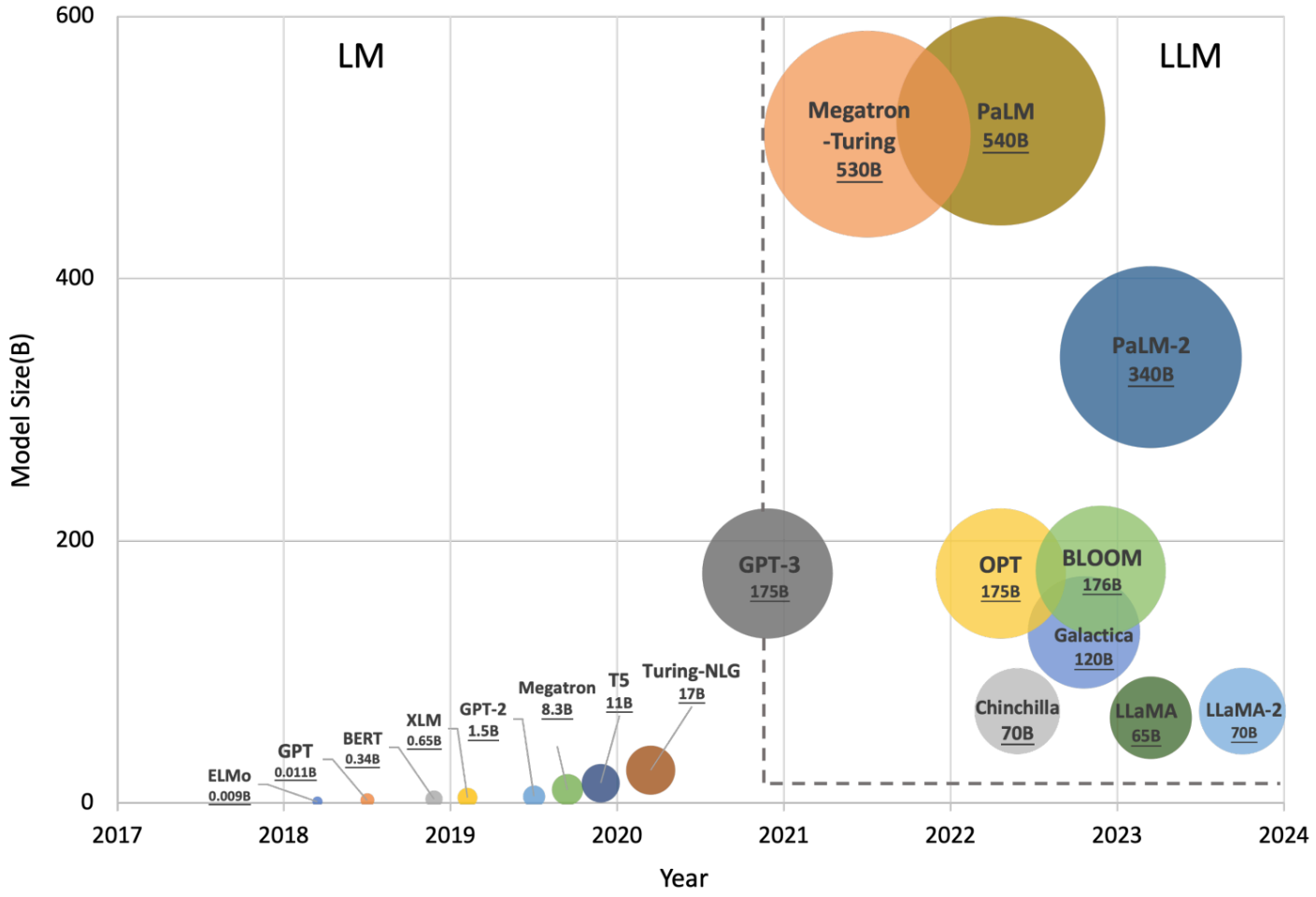
> ChatGPT is powered by a large language model (LLM) called **Generative Pre-trained Transformer** (GPT).

The LLM family



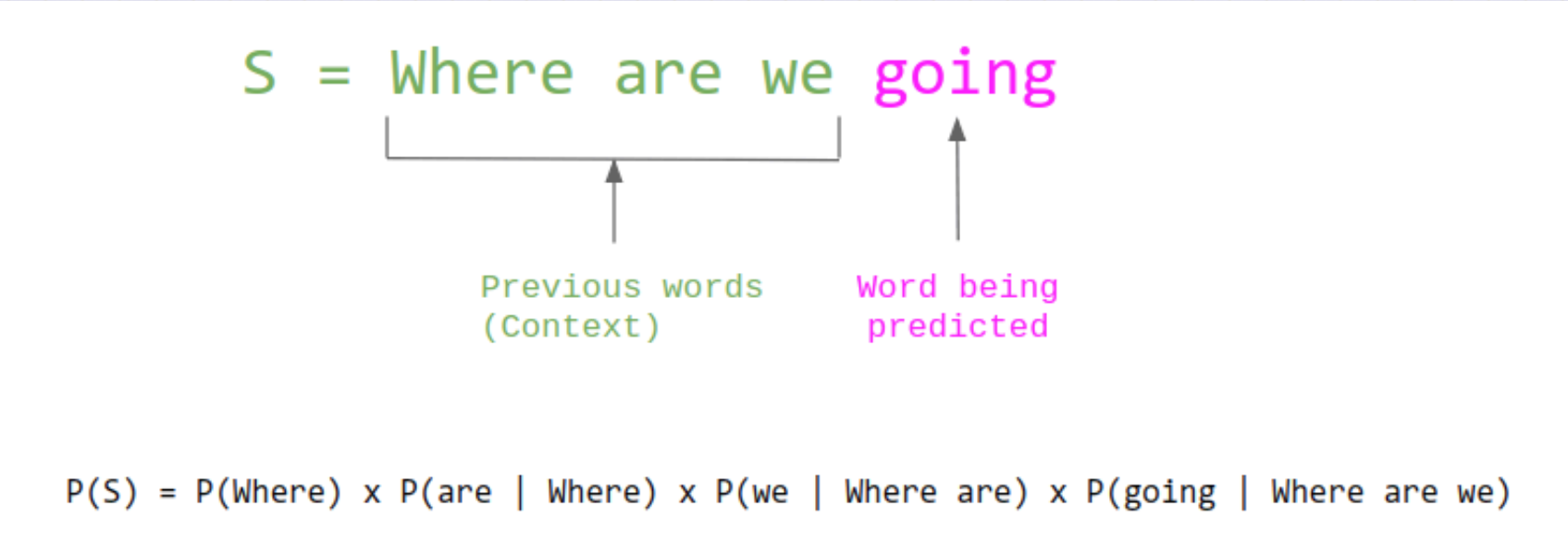
What makes LLM large?

> LLM is large because of the size of its **parameters**.



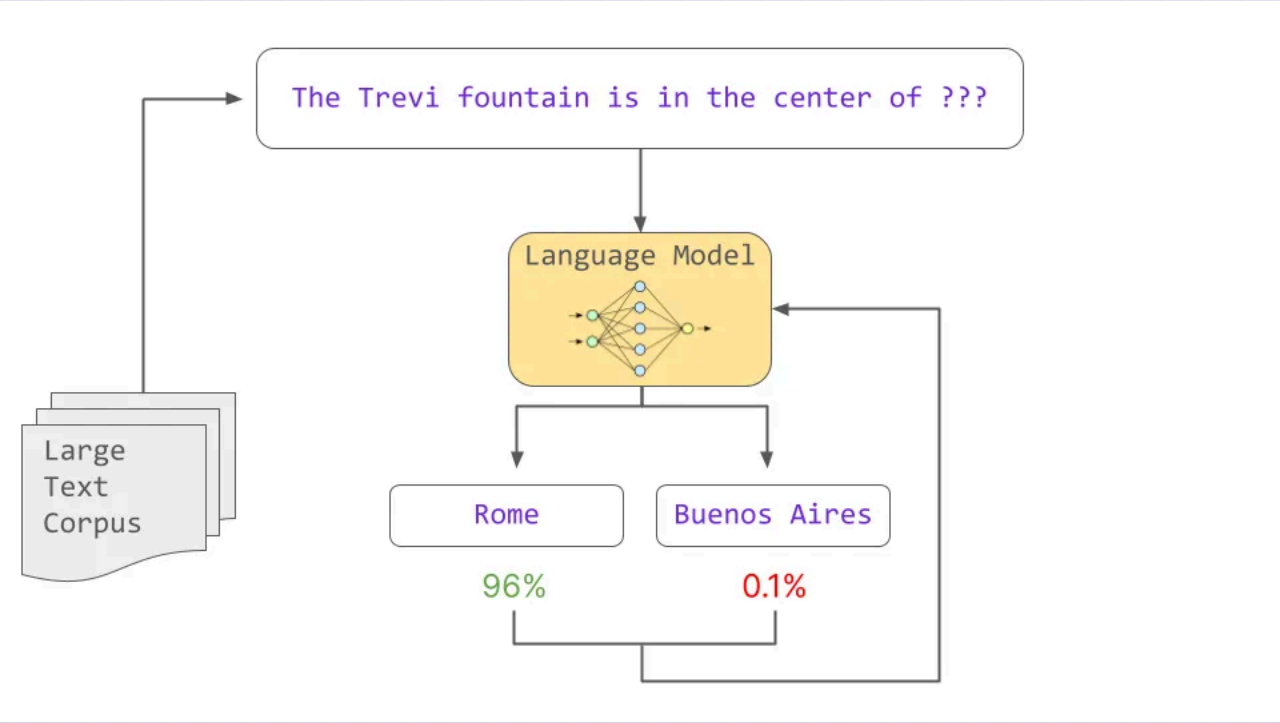
What is a language model?

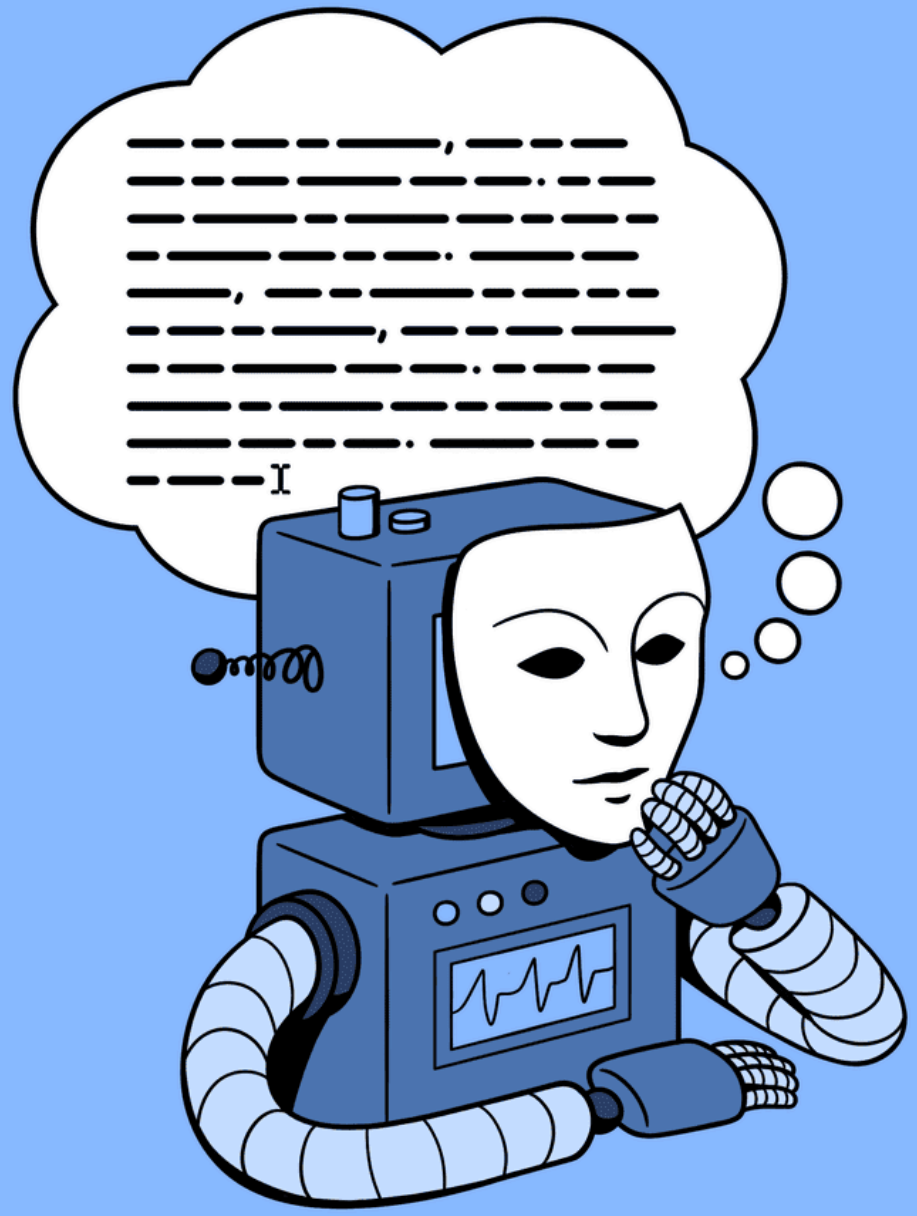
- A language model predicts the next word based on conditional probability.



A language model has some world knowledge.

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





Large Language Model (LLM)

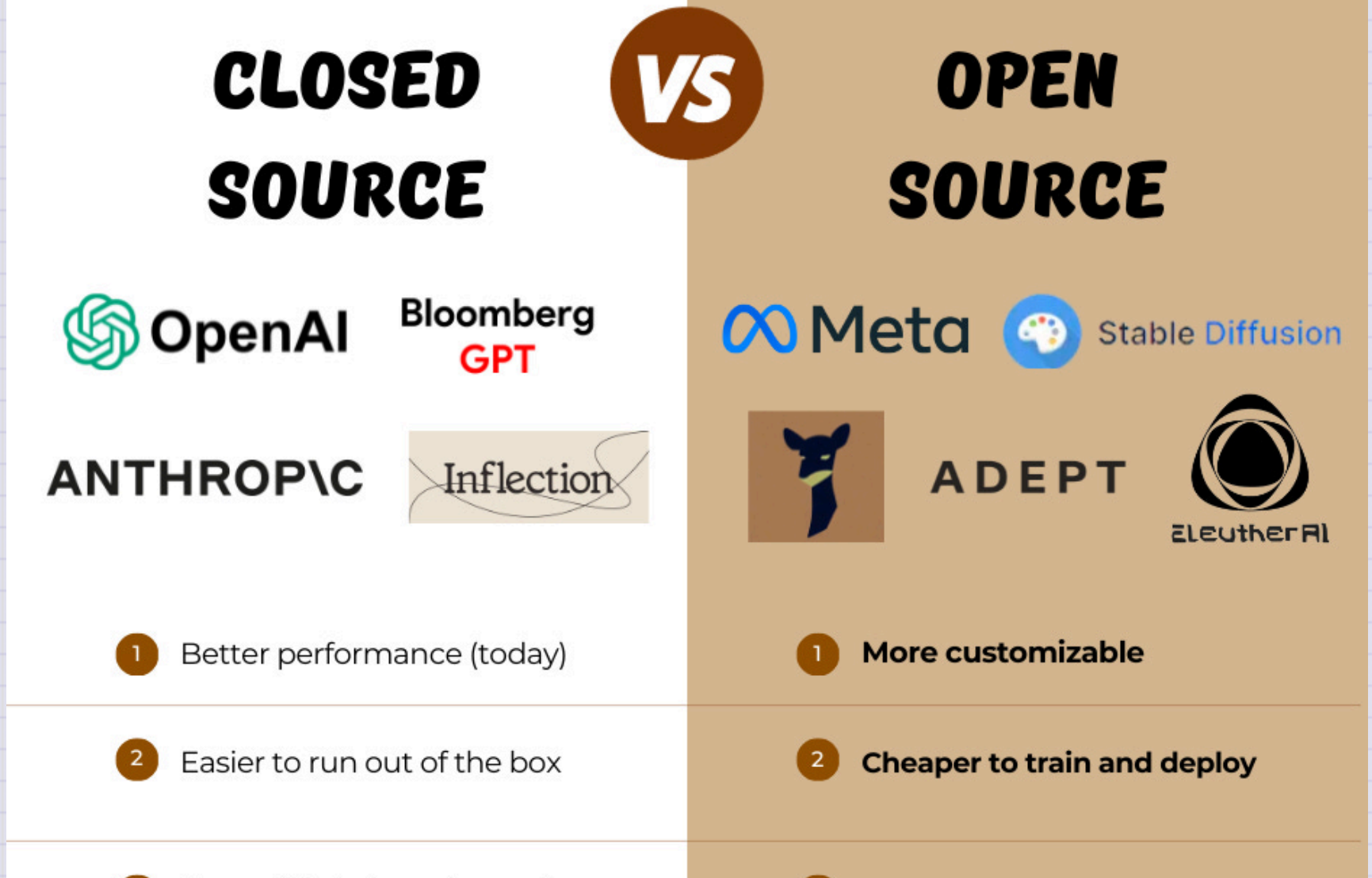
['lärj 'laŋ-gwij 'mä-dᵊl]

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

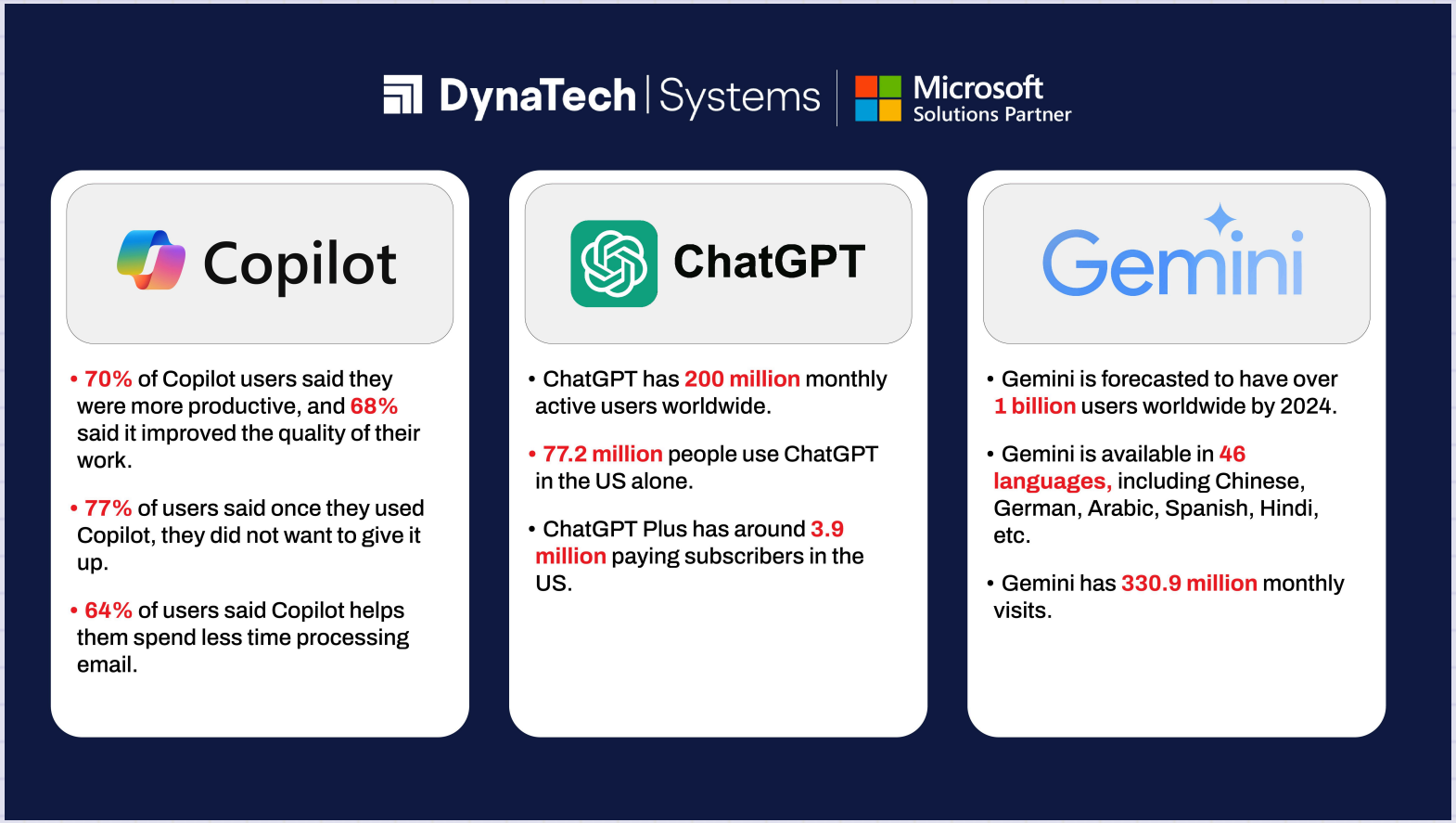
Investopedia

Closed-source LLMs

Open-source LLMs



Using closed-source LLMs



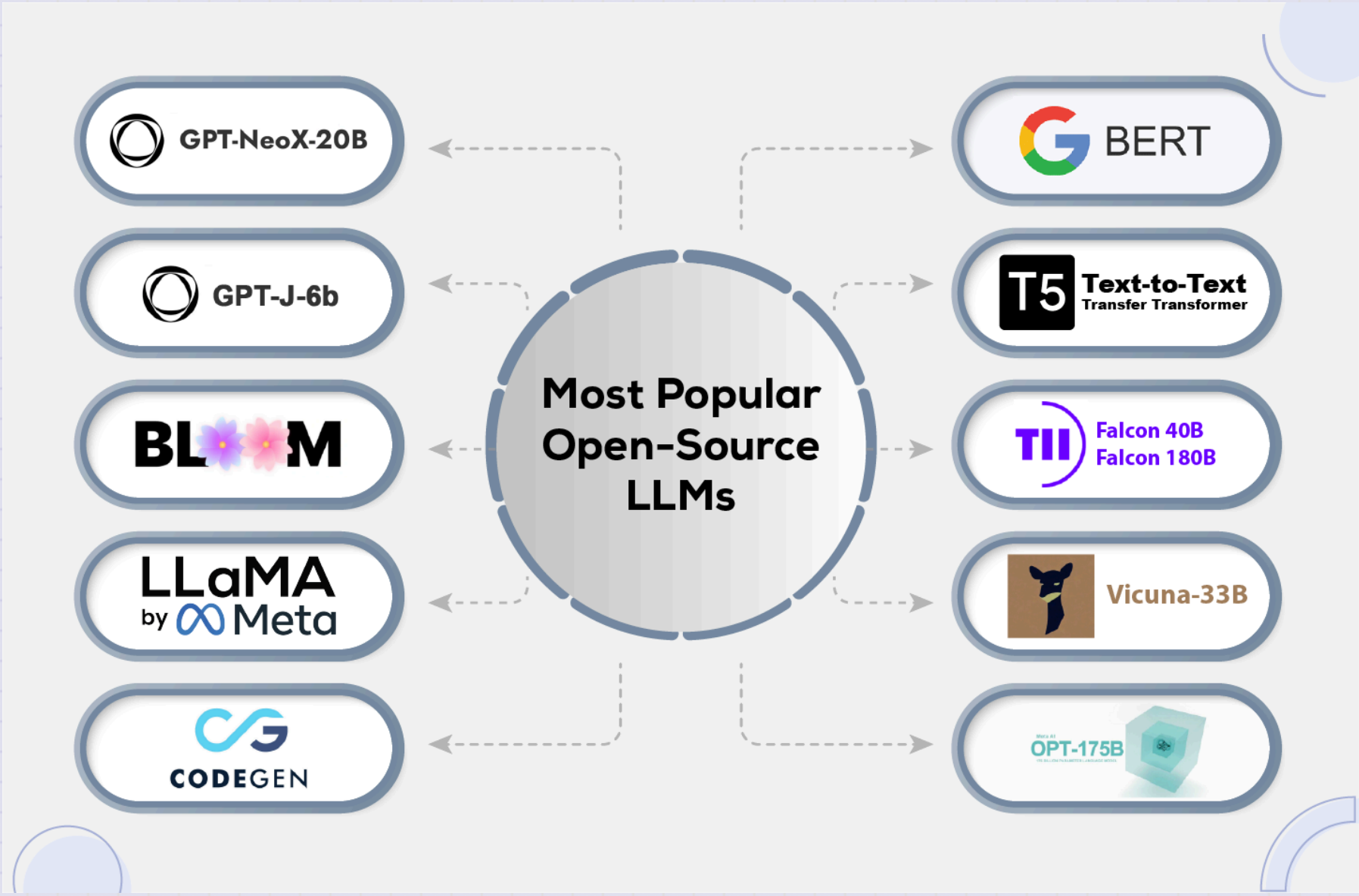
Using open-source LLMs

- gpt4all
- LM Studio
- Ollama

>> source

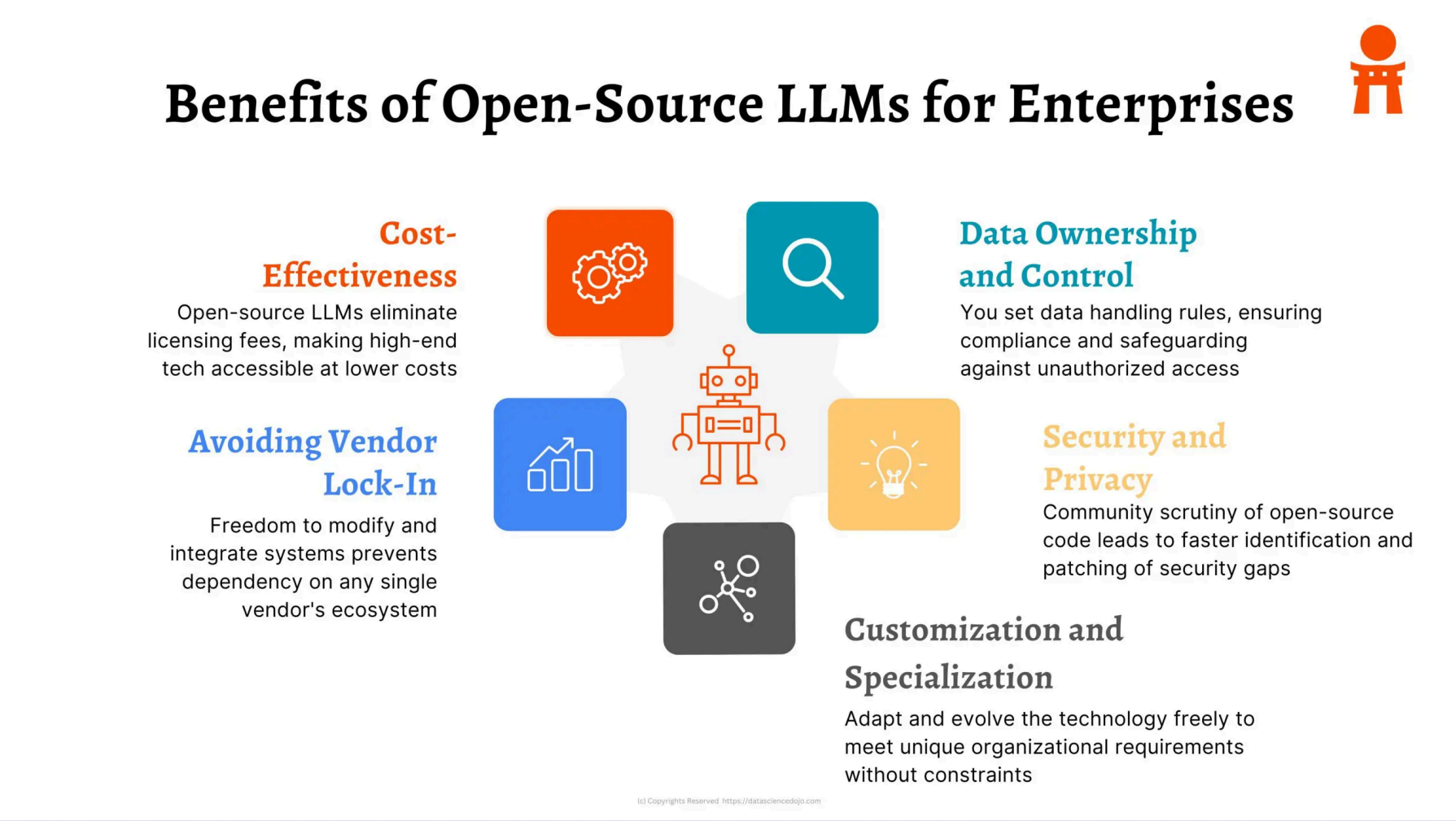
Popular open-source LLMs

>> source



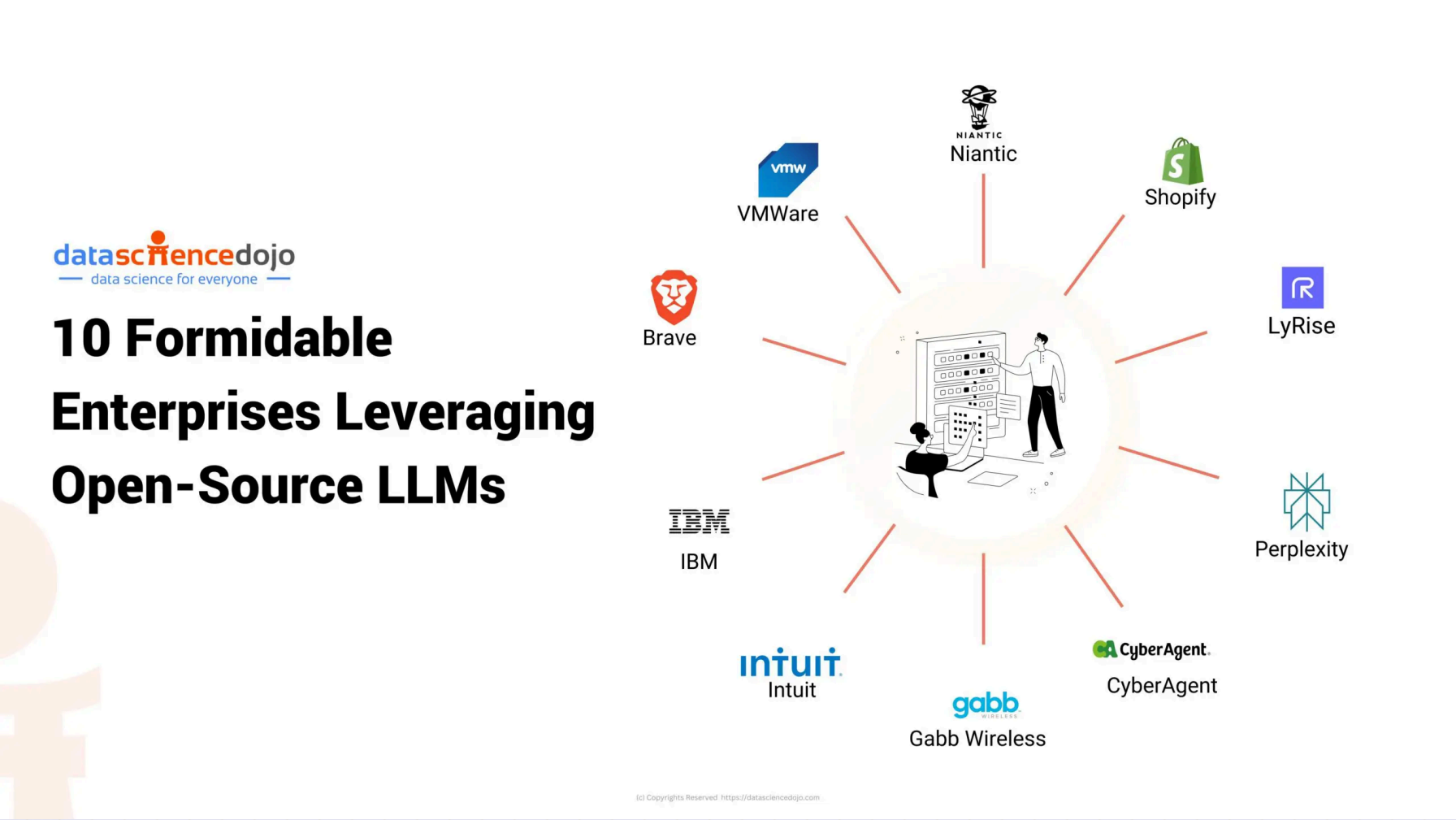
Benefits of open-source LLMs

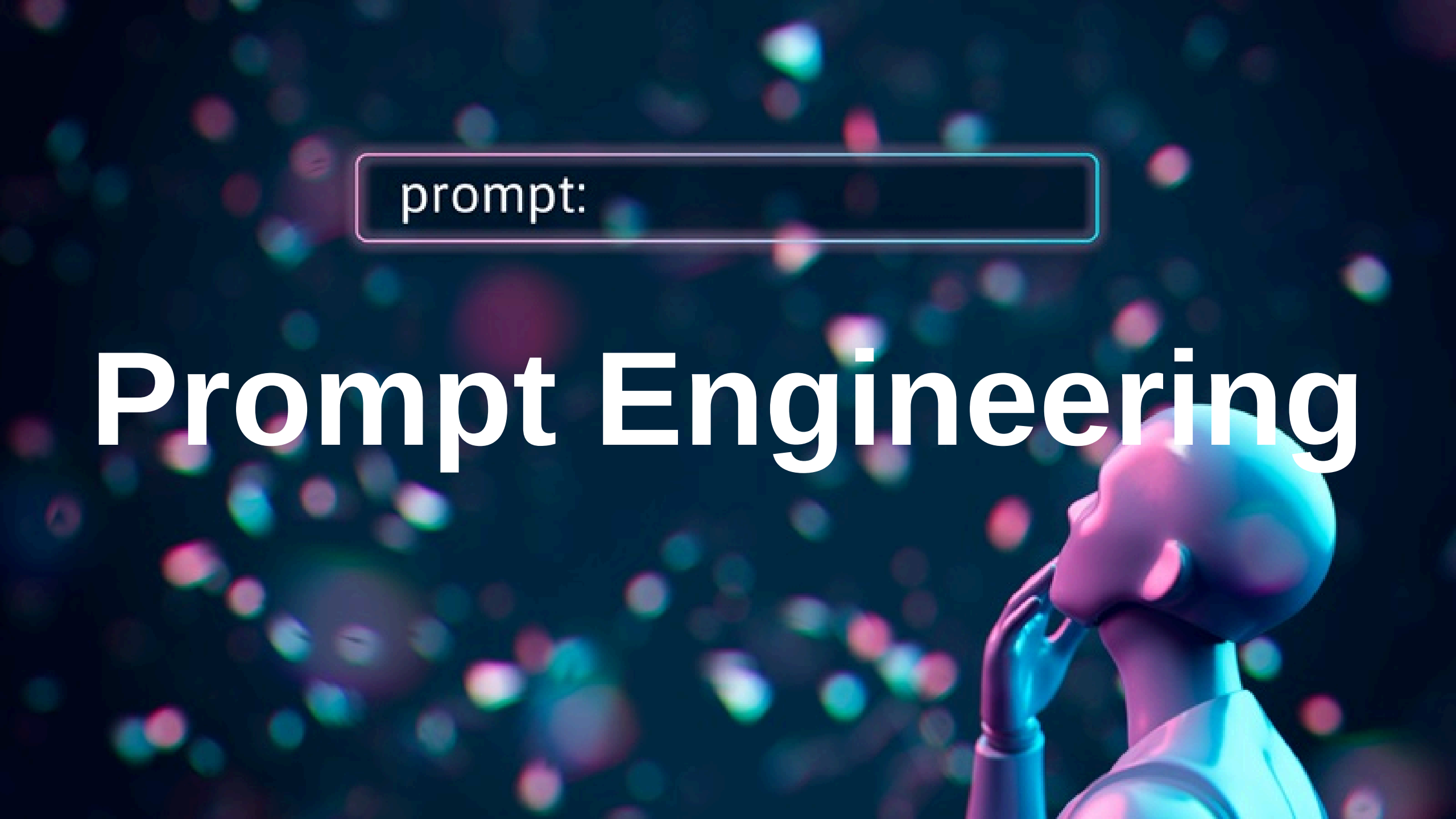
>> source



- IBM
- Perplexity

>> source



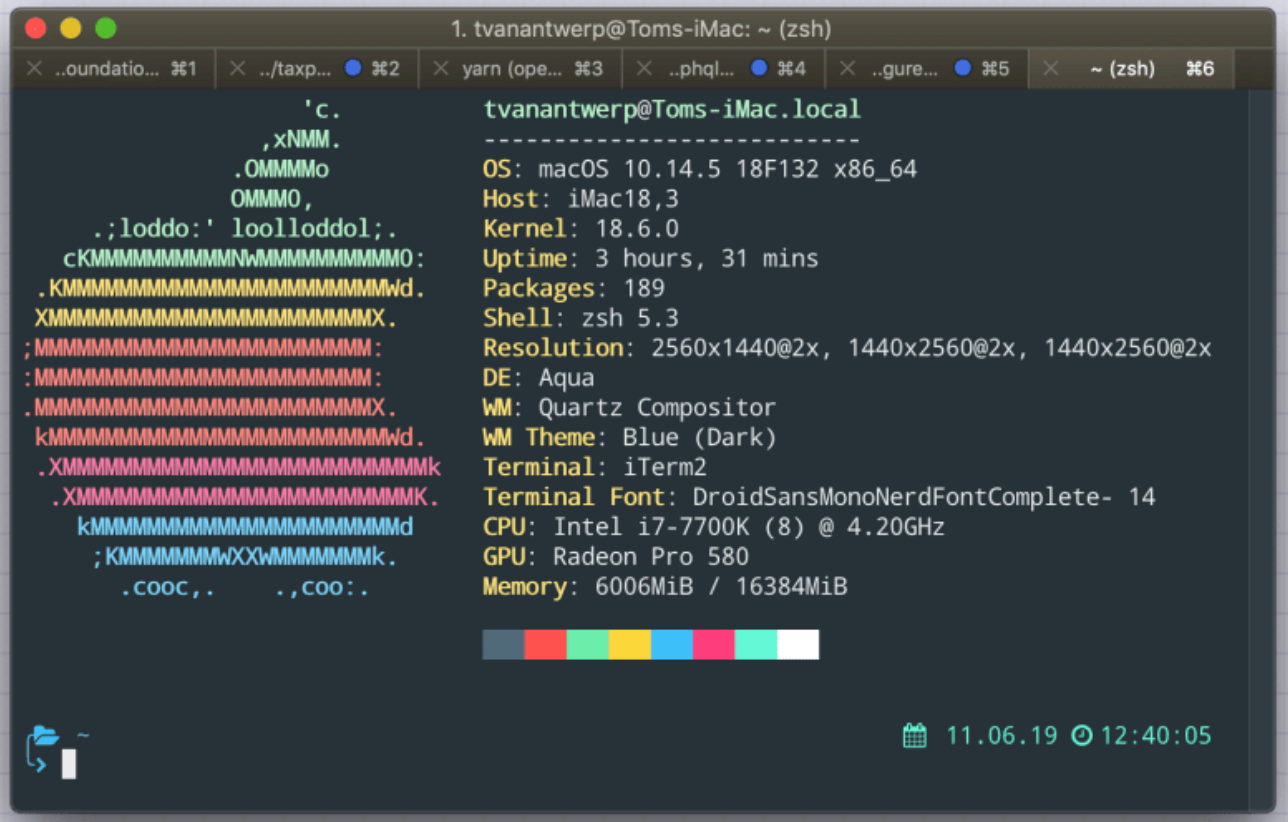


Prompt Engineering

Q1: How do you design prompts to effectively elicit the desired responses?

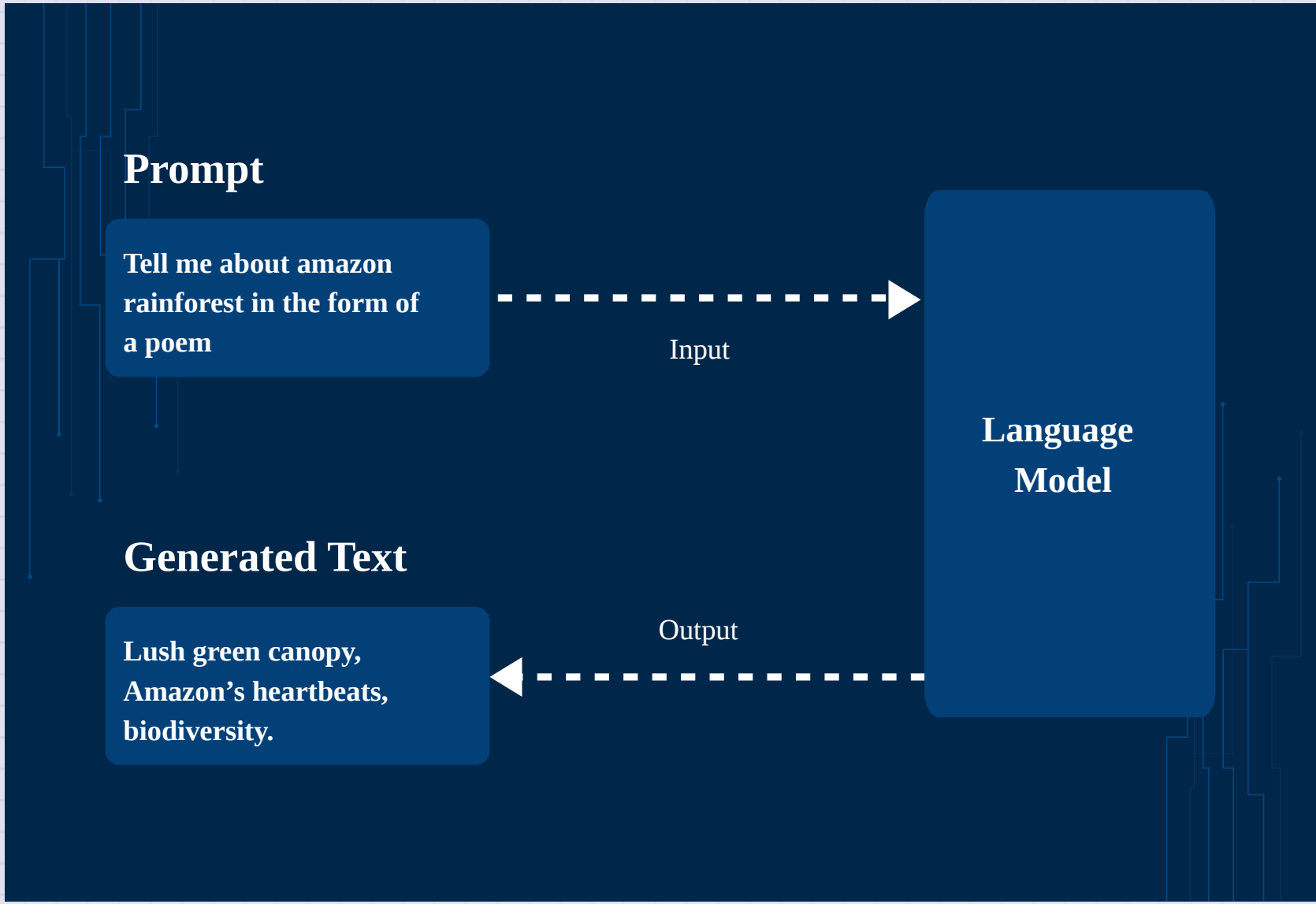
Q2: How can humans and AI work together to create more effective and creative outcomes?

Prompt on the terminal

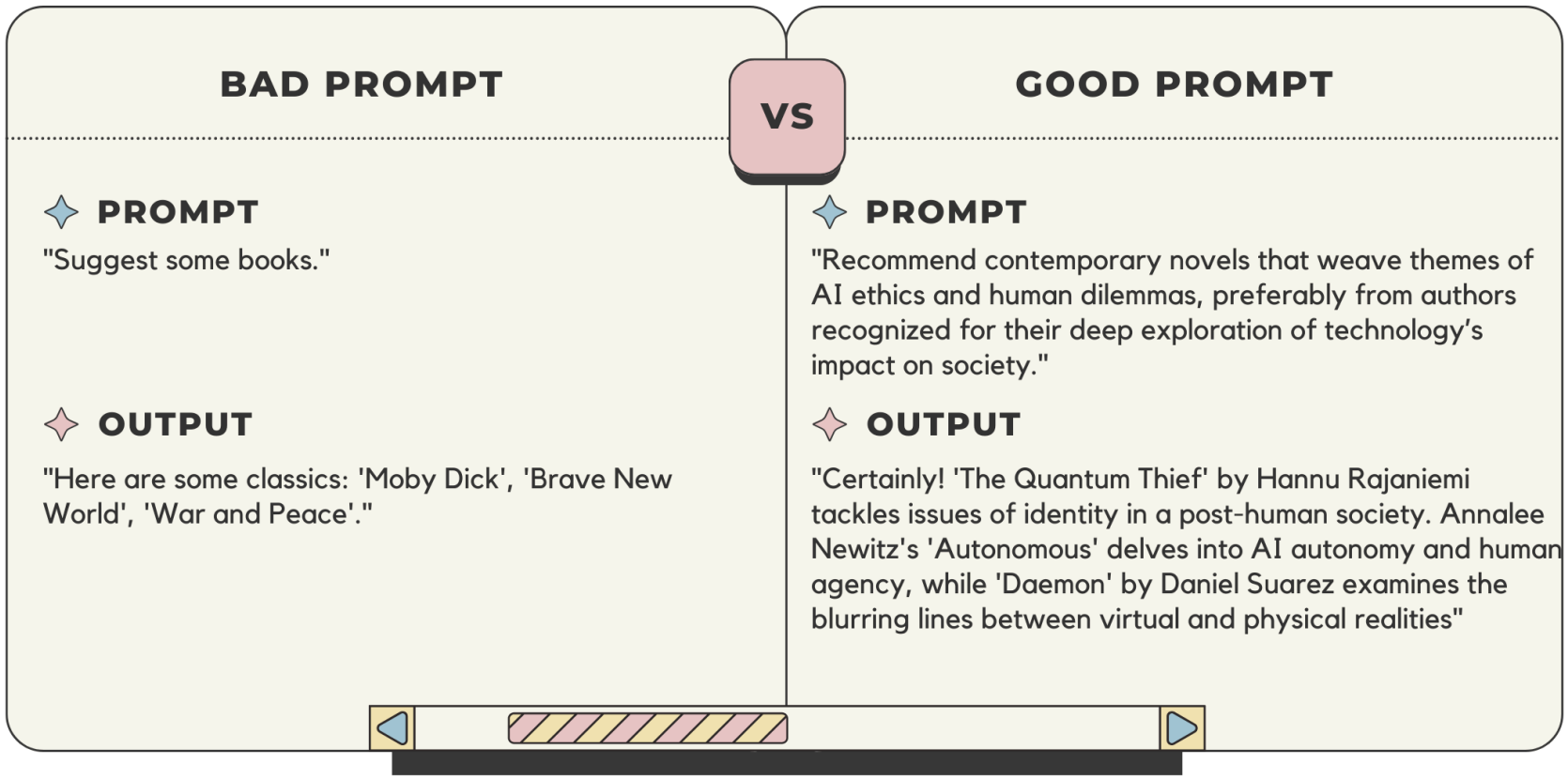


Prompt in GenAI

- 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

Checklist

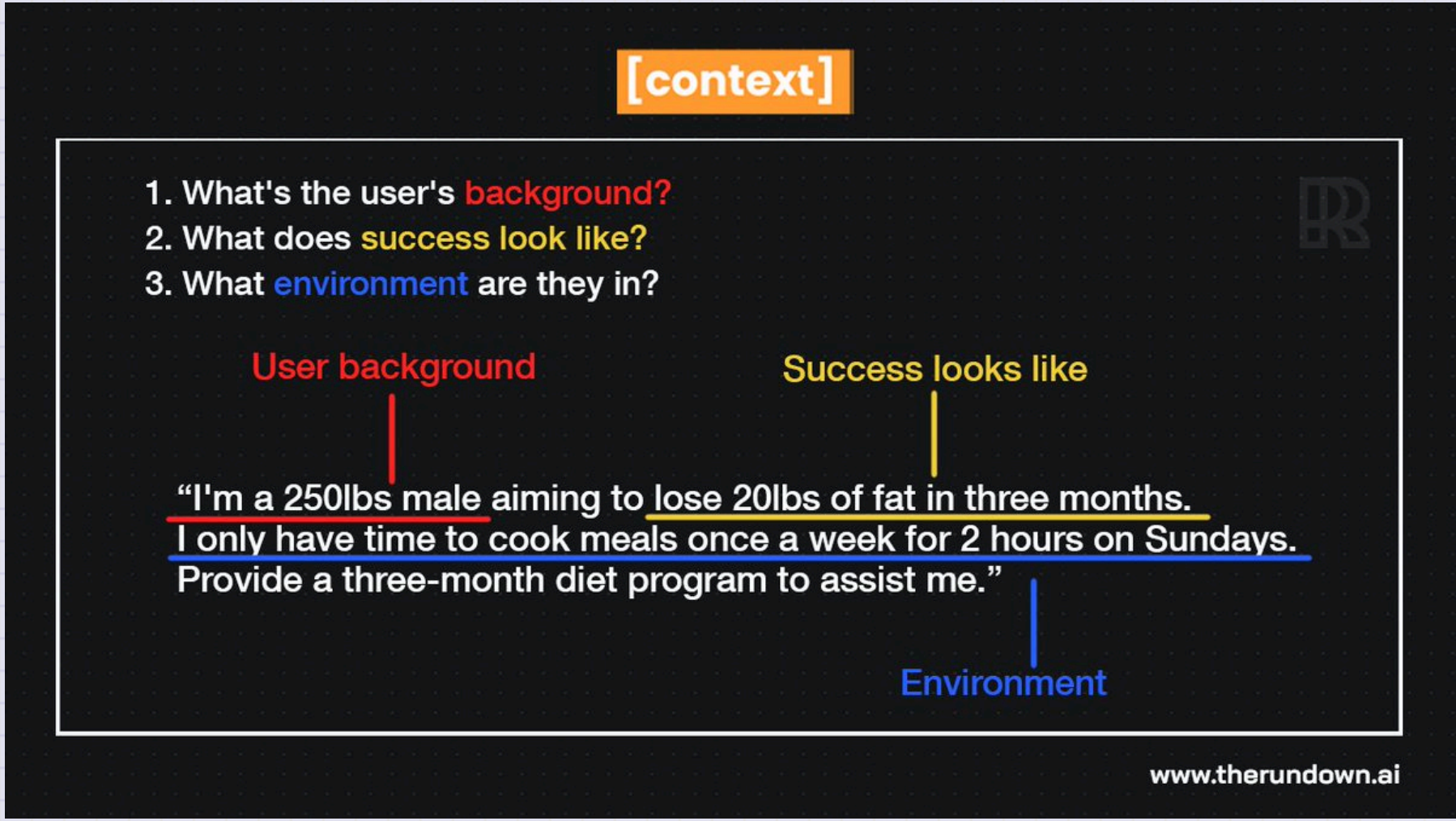
>> source

01: Task

>> source

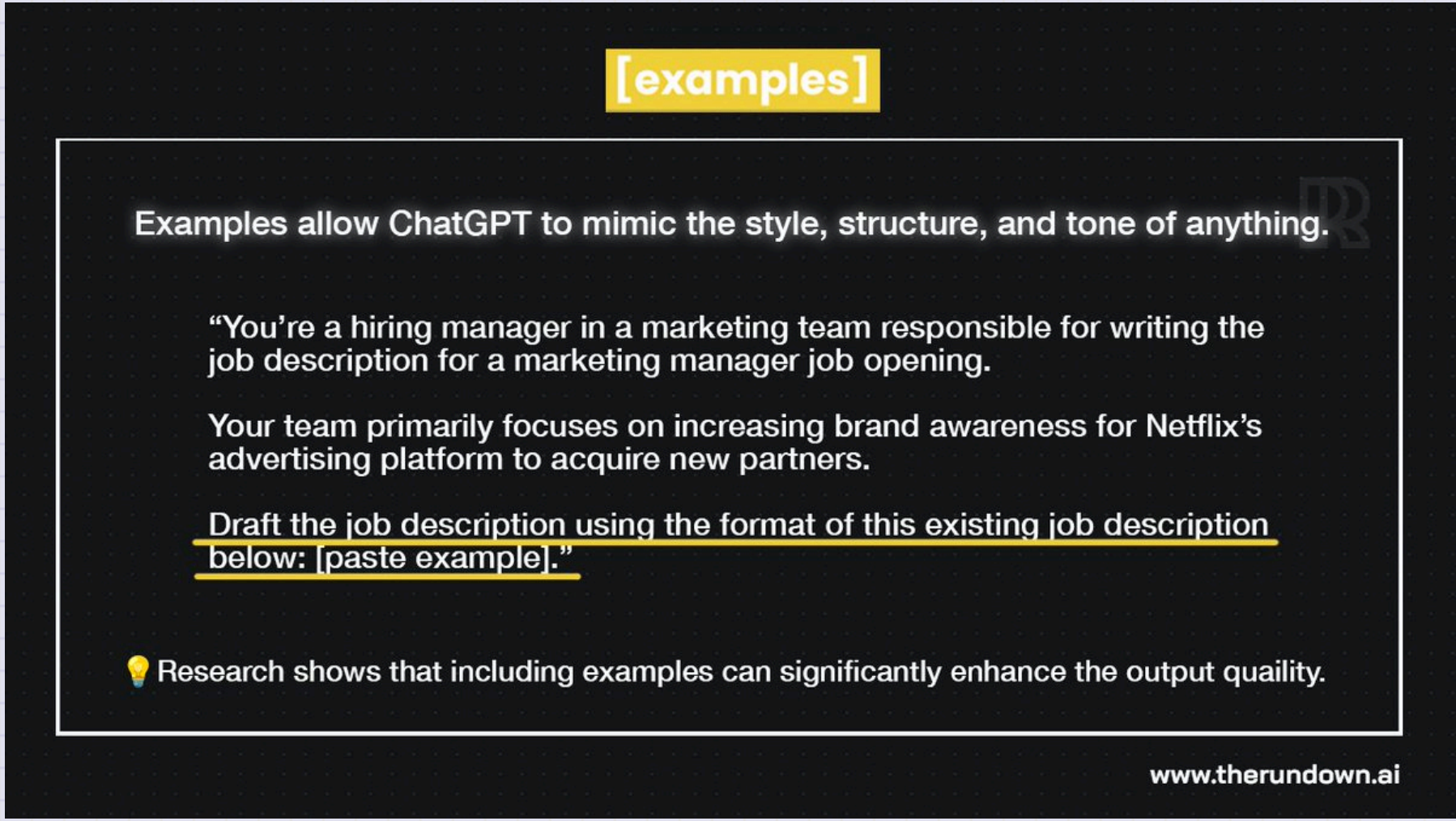
02: Context

>> source



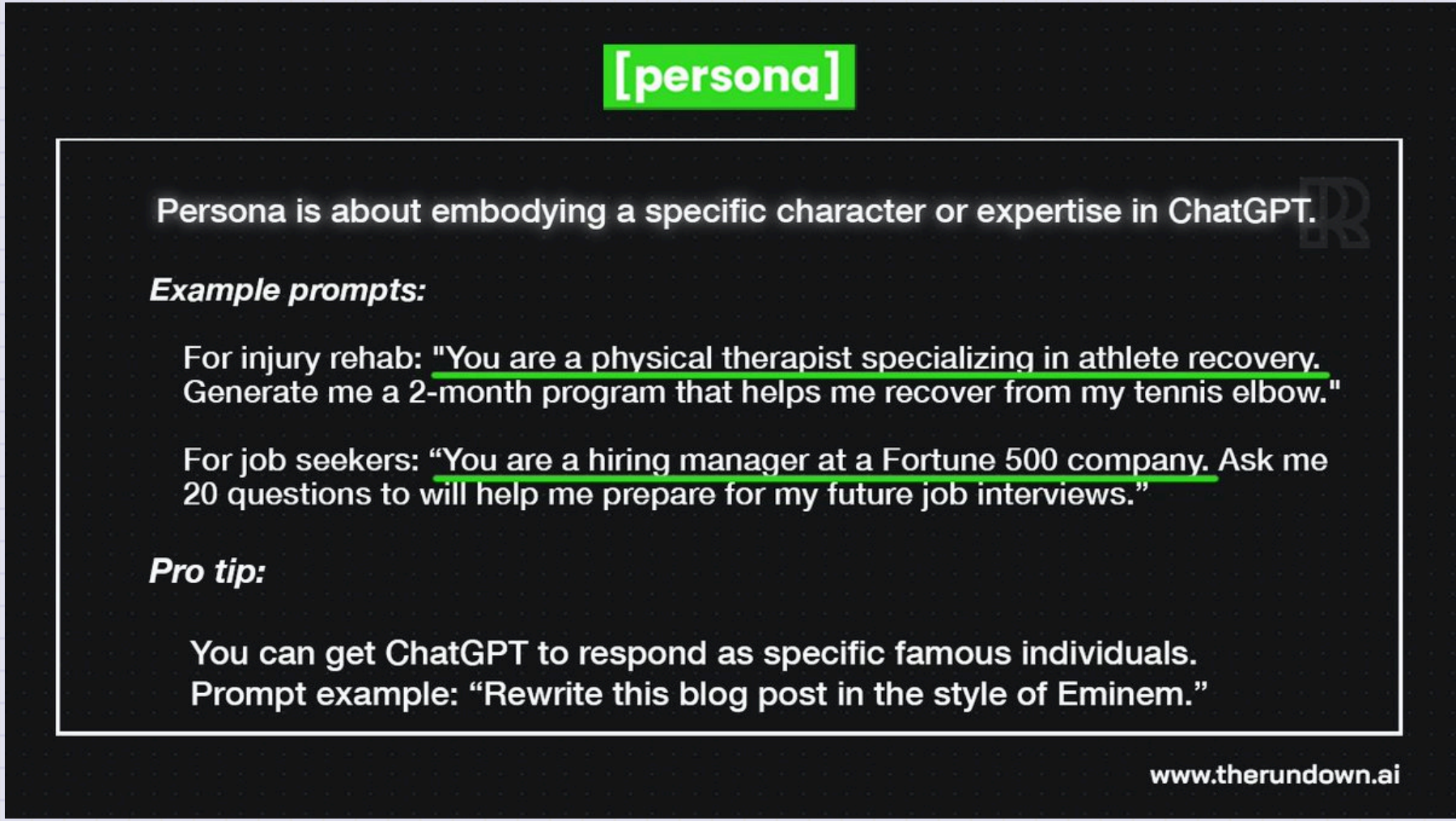
03: Examples

>> source



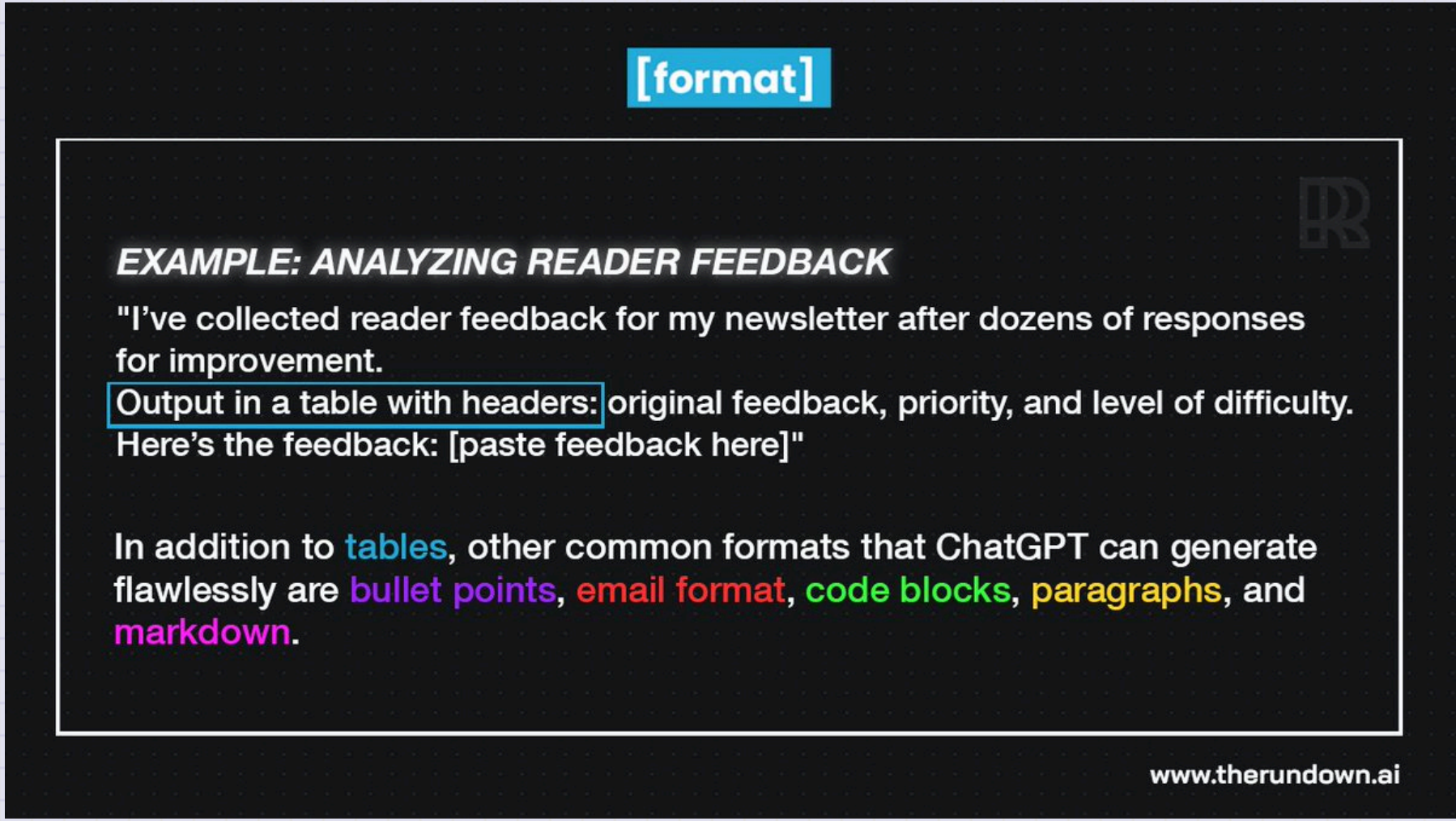
04: Persona

>> source



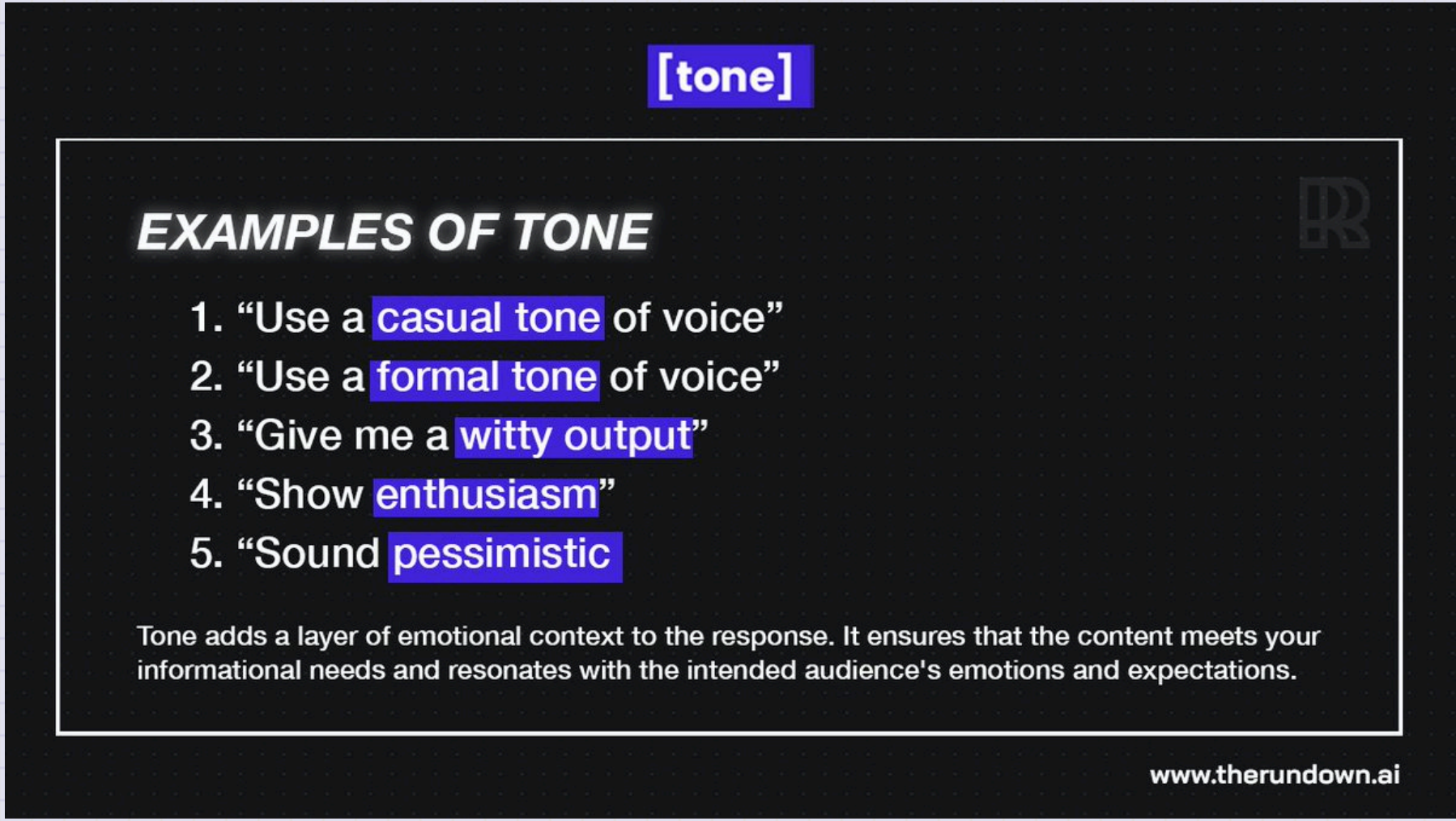
05: Format

>> source



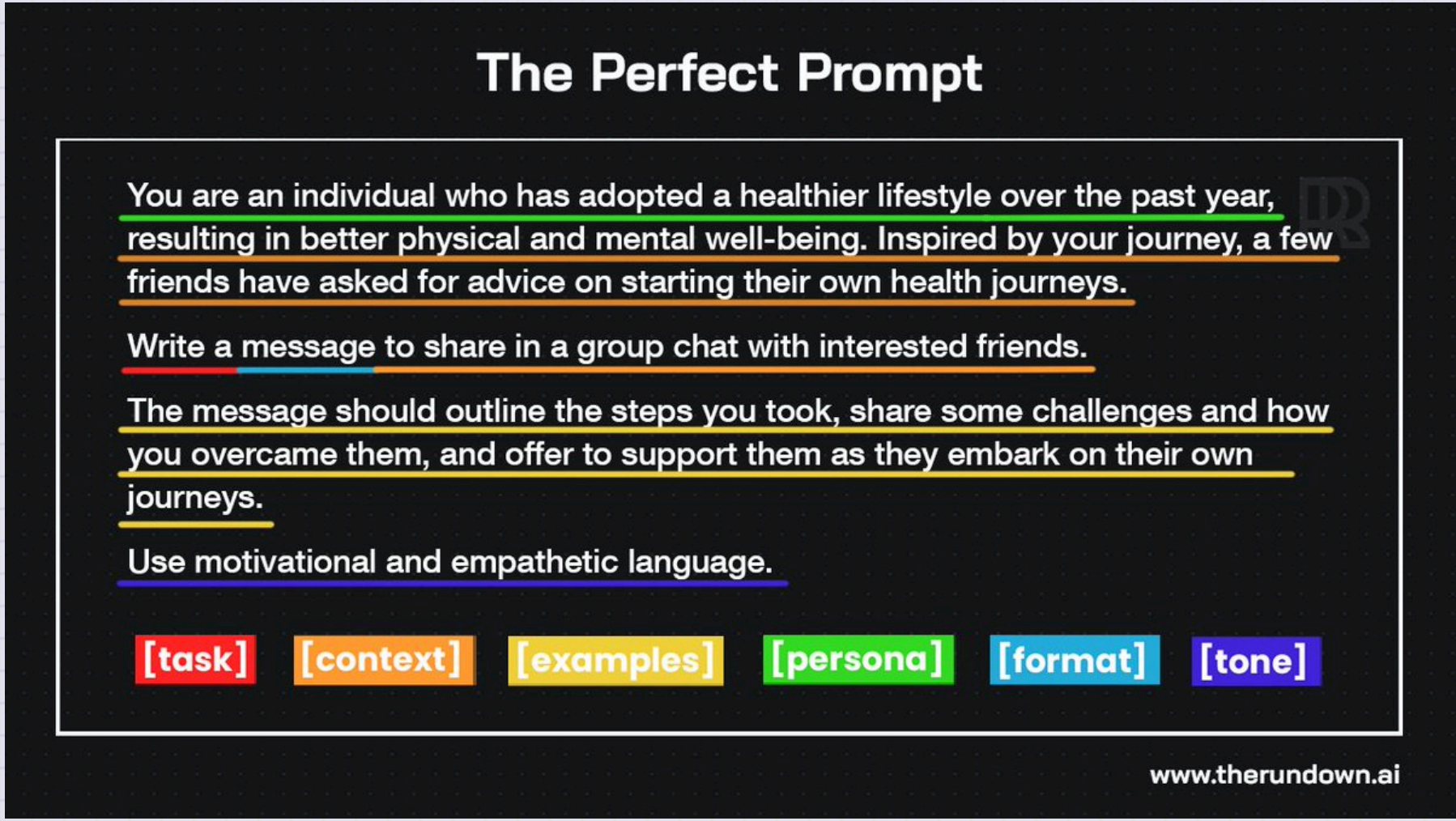
06: Tone

>> source



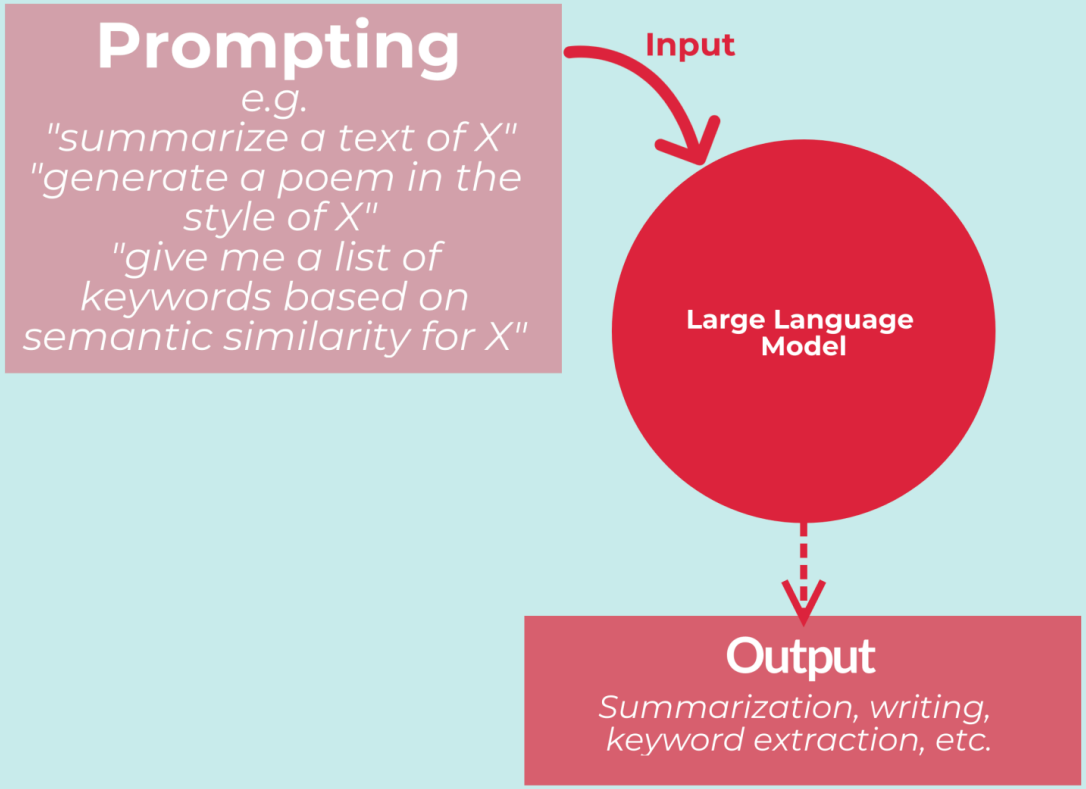
Perfect prompt

>> source



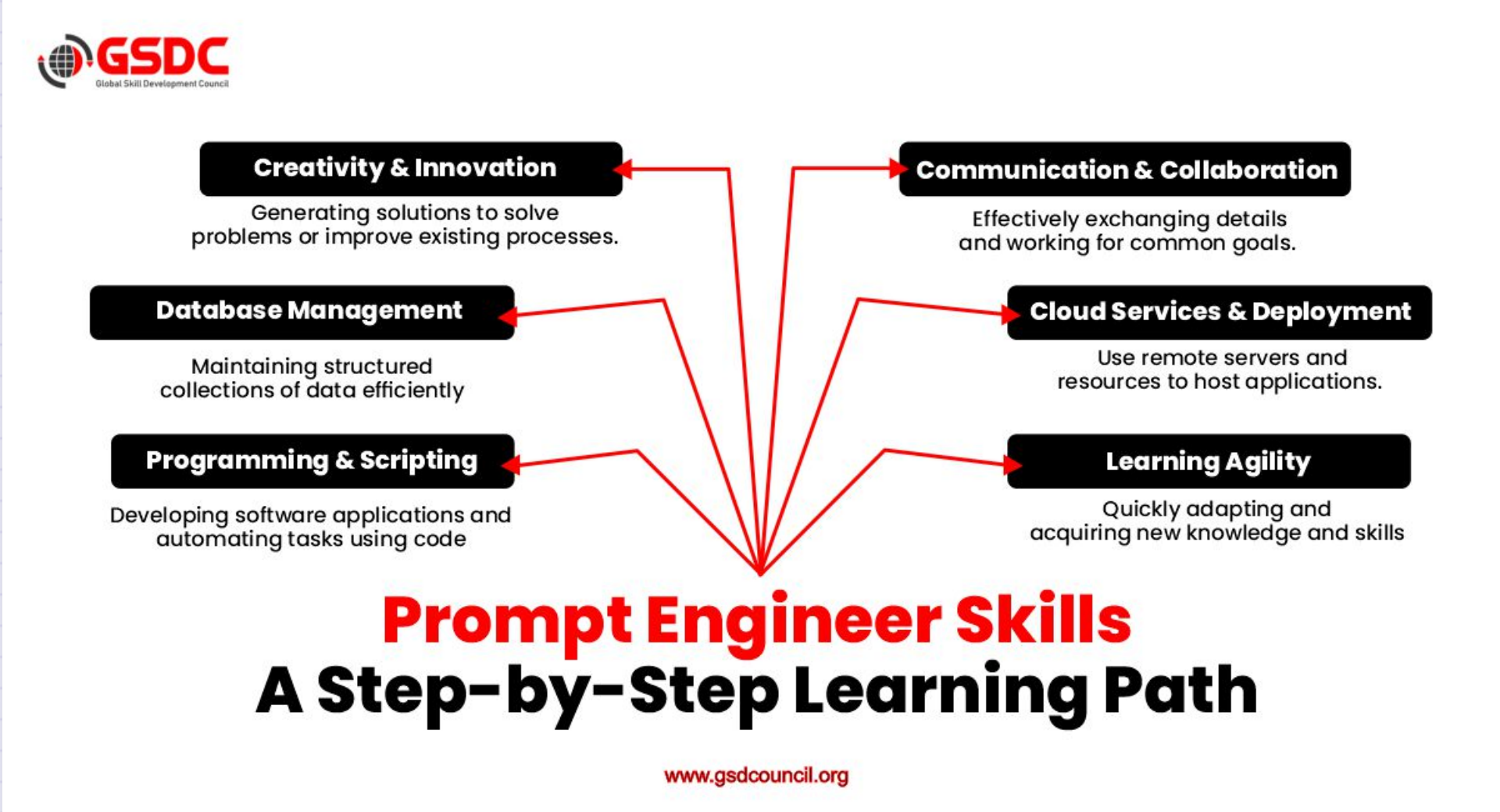
Prompt Engineering In A Nutshell

- Prompt engineering is a natural language processing (NLP) concept that involves discovering inputs that yield desirable or useful results.
- Like most processes, the quality of the inputs determines the quality of the outputs in prompt engineering. Designing effective prompts increases the likelihood that the model will return a response that is both favorable and contextual.
- Developed by OpenAI, the CLIP (Contrastive Language-Image Pre-training) model is an example of a model that utilizes prompts to classify images and captions from over 400 million image-caption pairs.



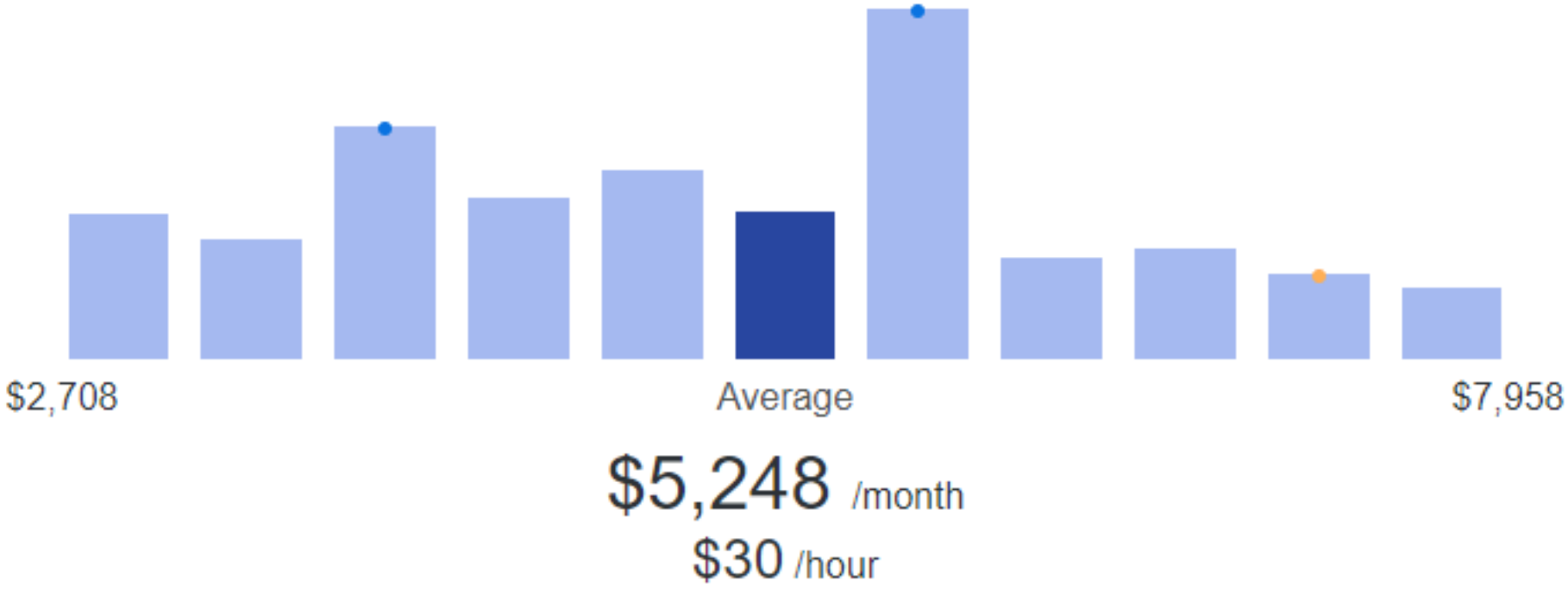
Skills

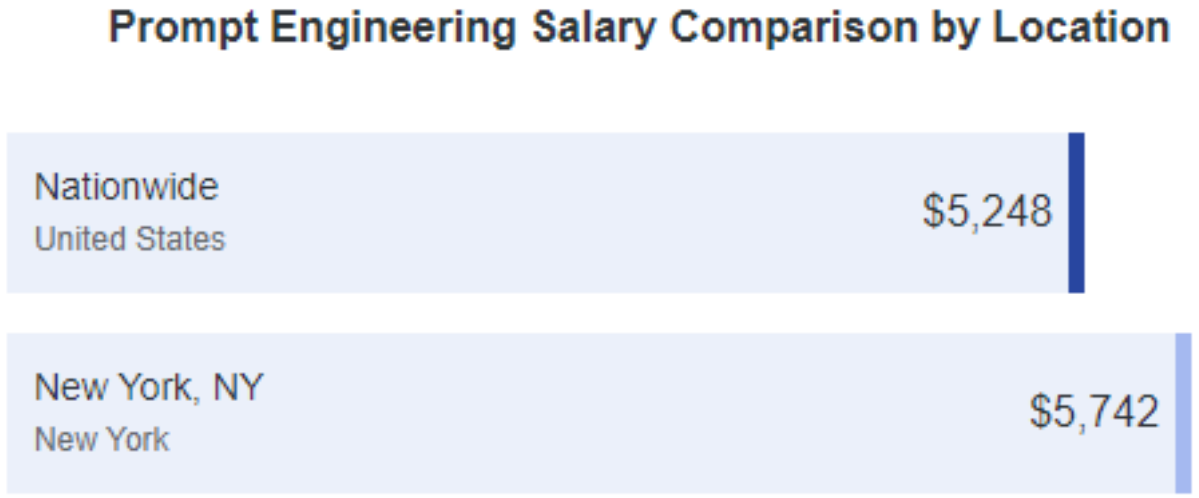
>> source



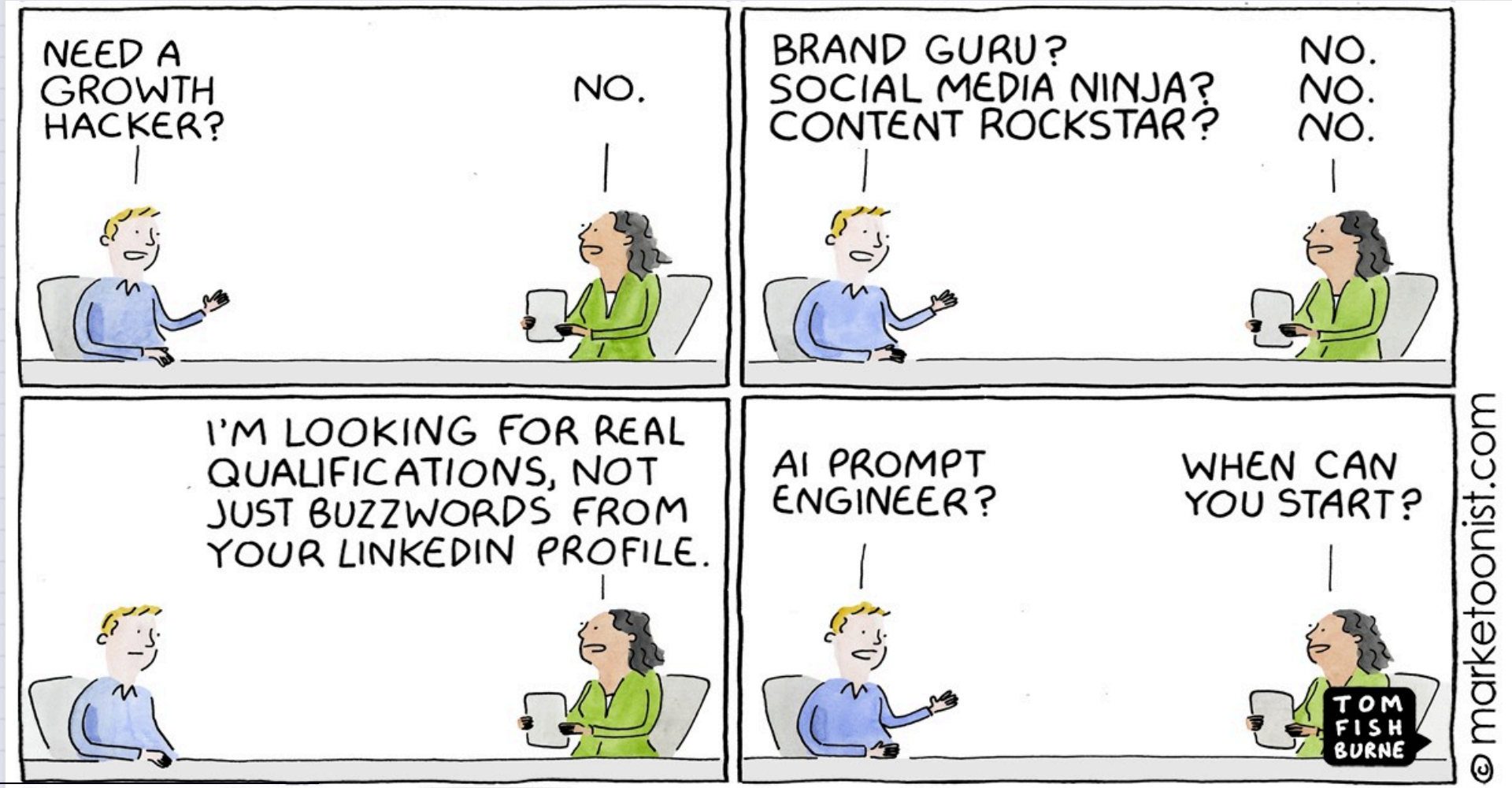
Salary

>> source

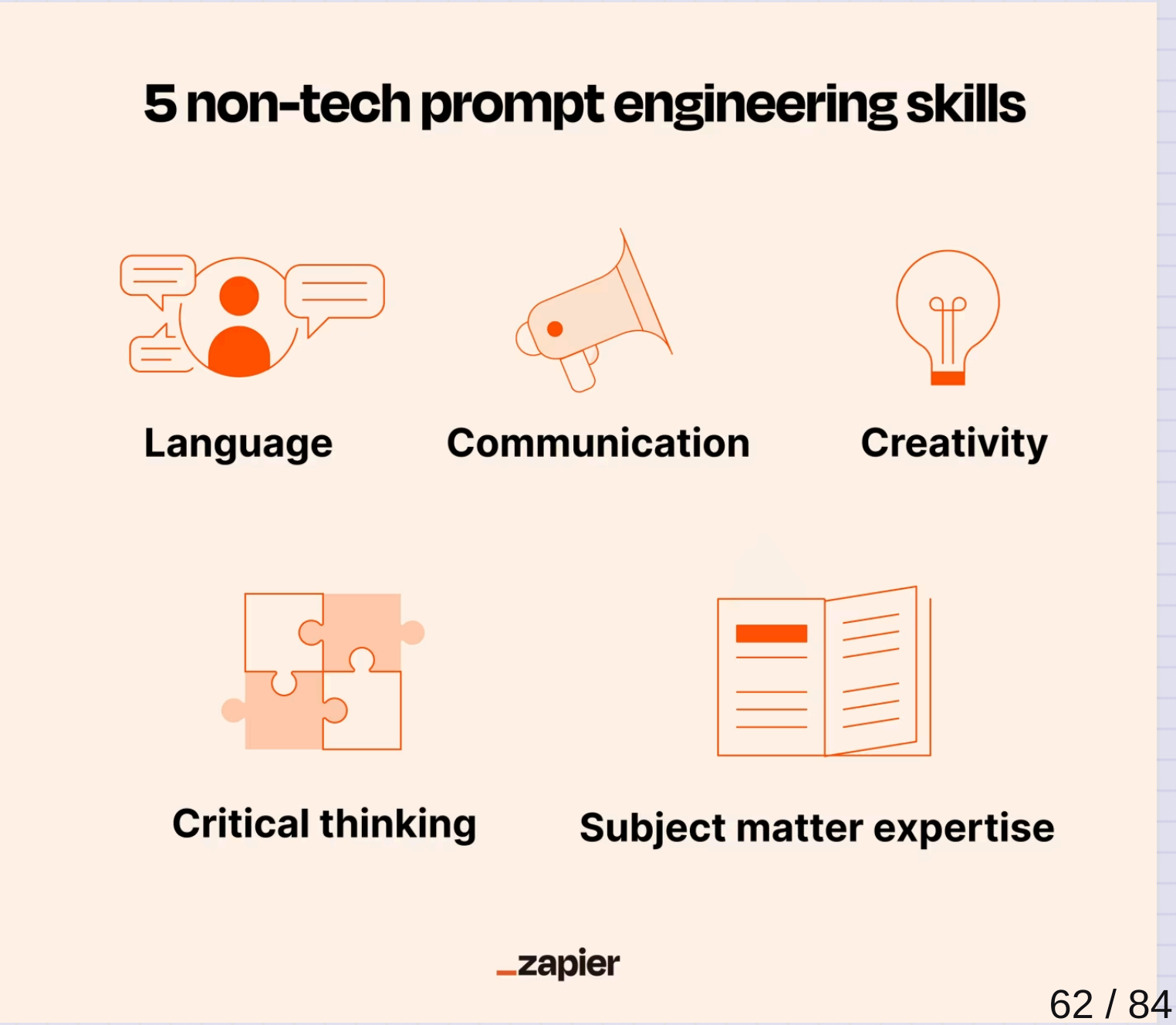




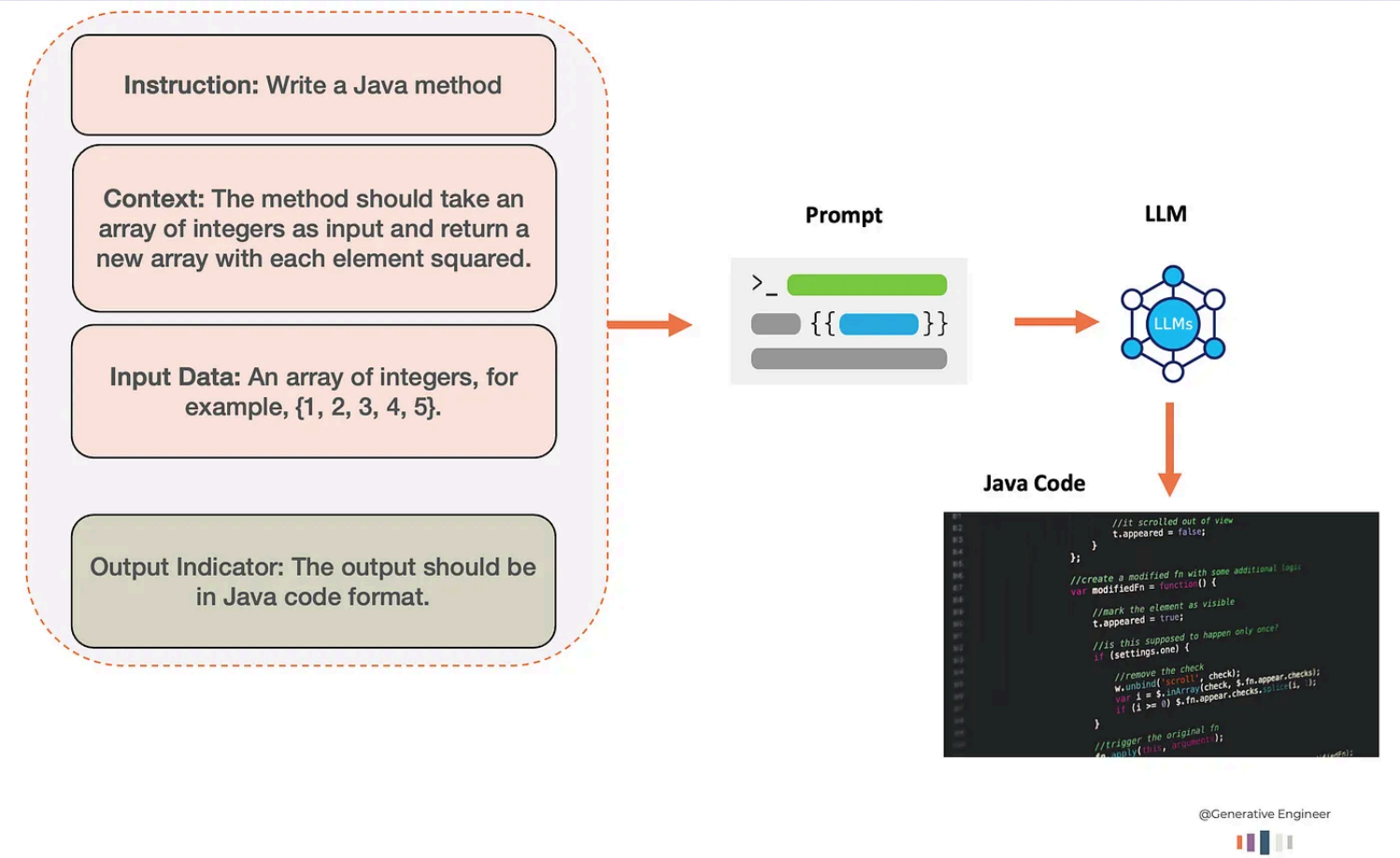
Prompt engineers are in high demand.



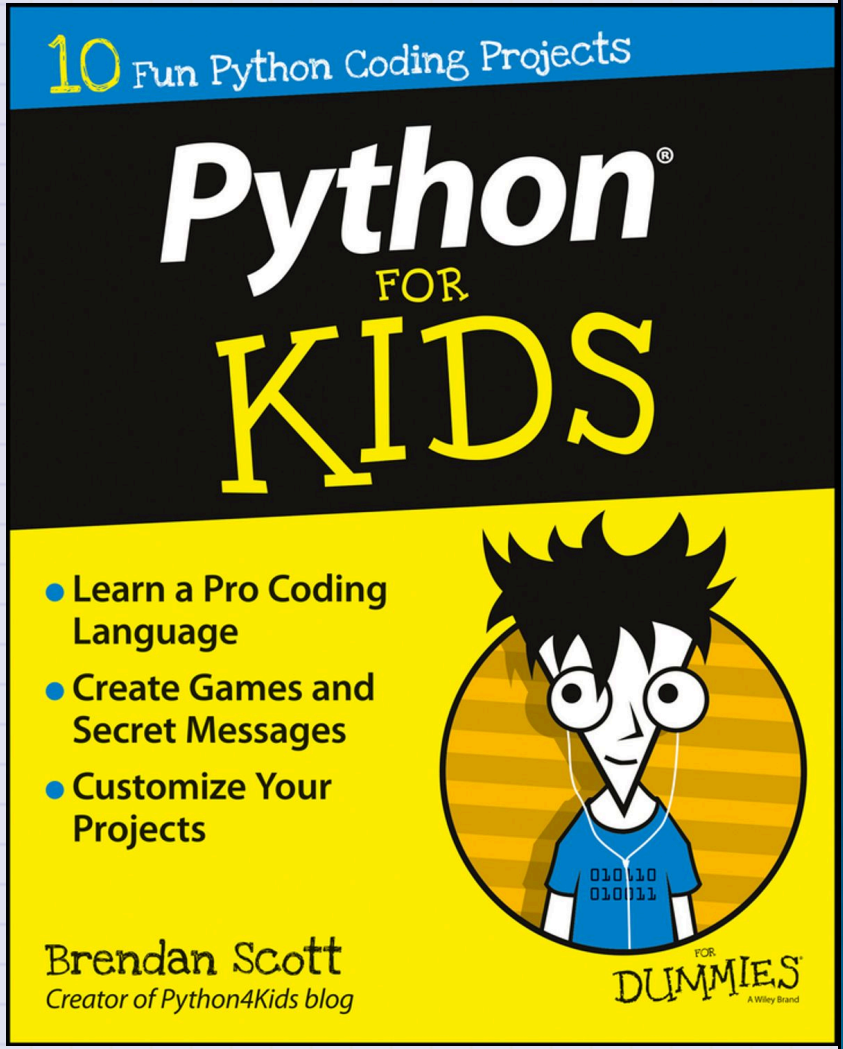
Soft skills for prompt engineering

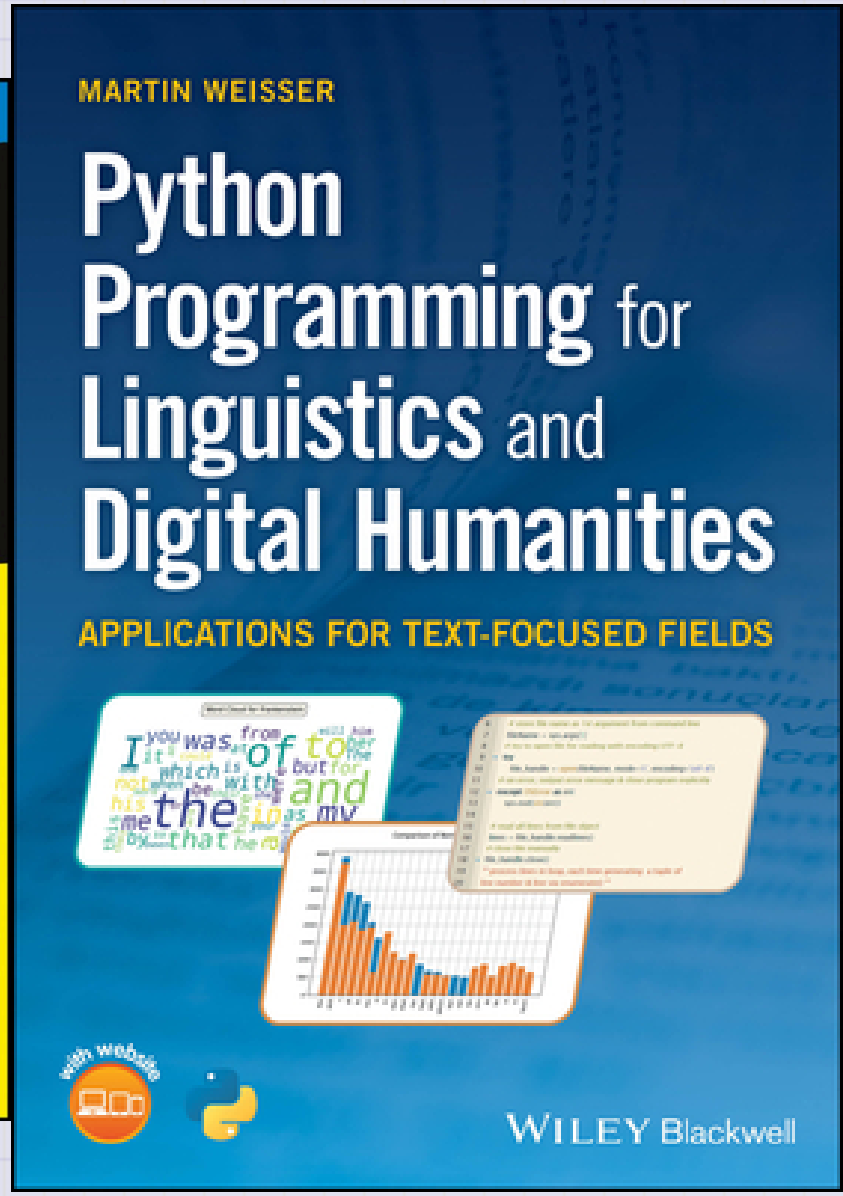


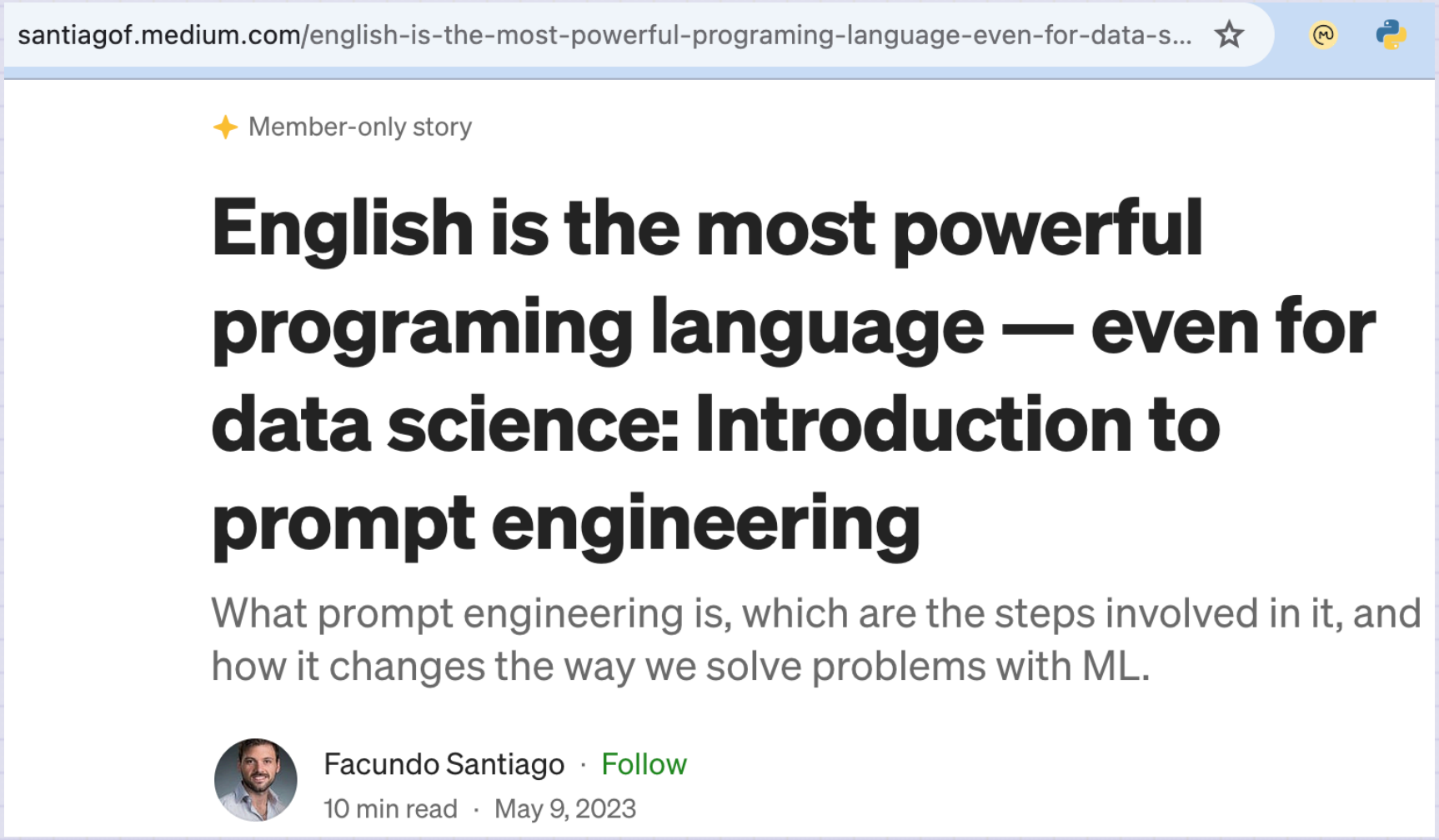
Technical skills for prompt engineering



Start with Python if you want to try programming.






santiagof.medium.com/english-is-the-most-powerful-programing-language-even-for-data-s...
Member-only story
English is the most powerful programing 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

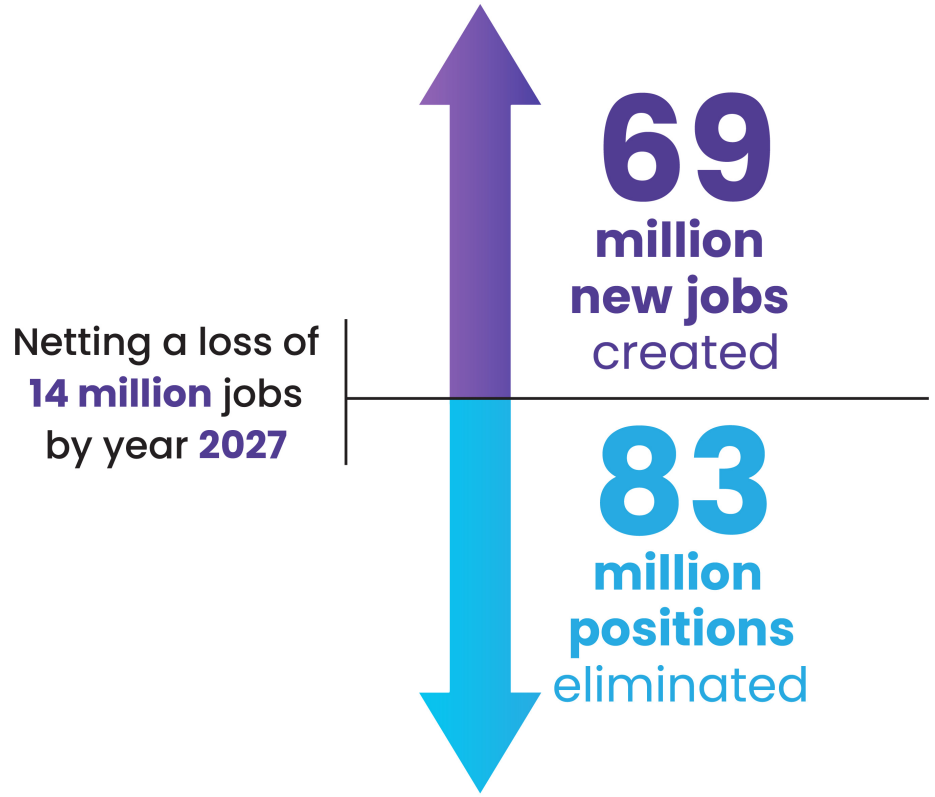
>> source

With LLMs, everyone can be a developer, even if you're just 9 years old.

>> source

GenAI's impact on future jobs

>> source

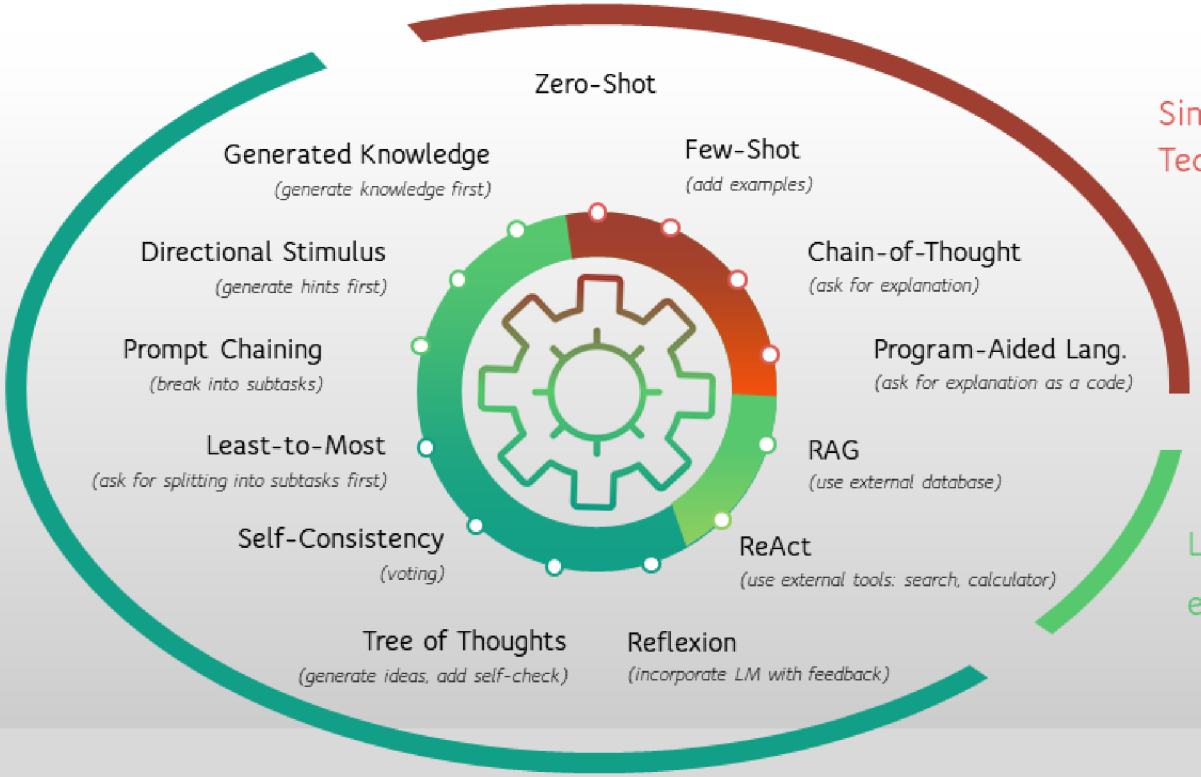


69 million new jobs created and **83 million** positions eliminated by **2027,** netting a loss of **14 million** jobs.

(Source: World Economic Forum)

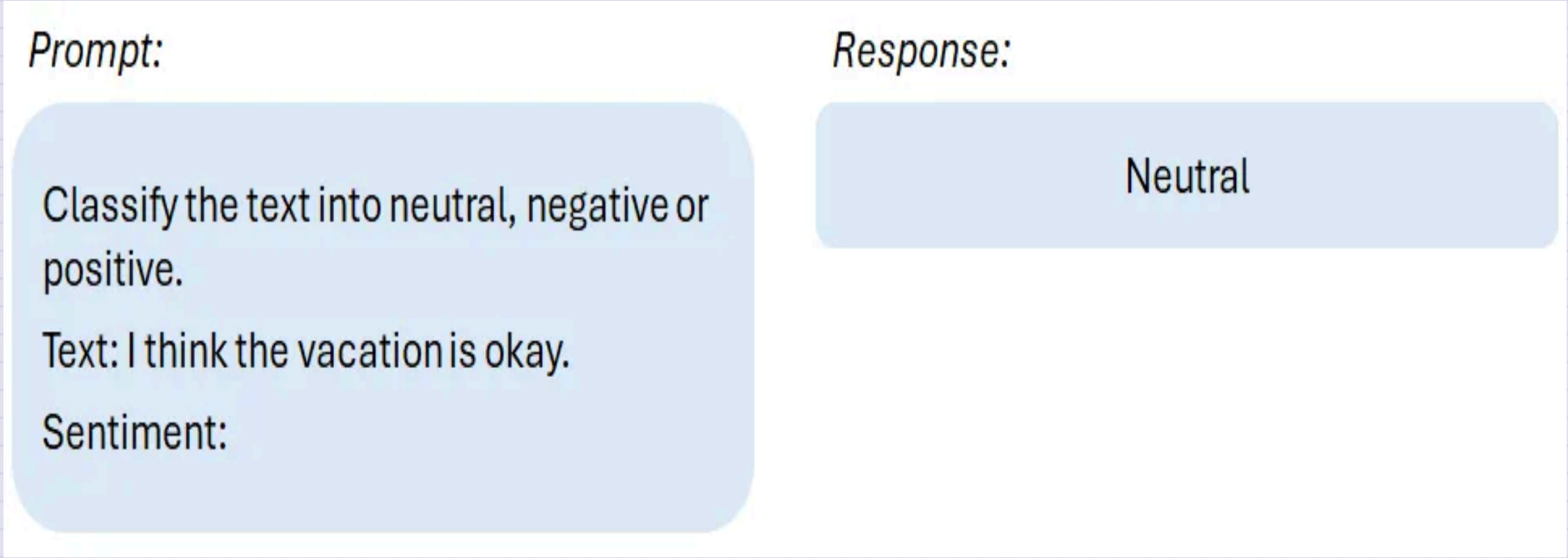
Prompting techniques

Prompt Engineering Techniques



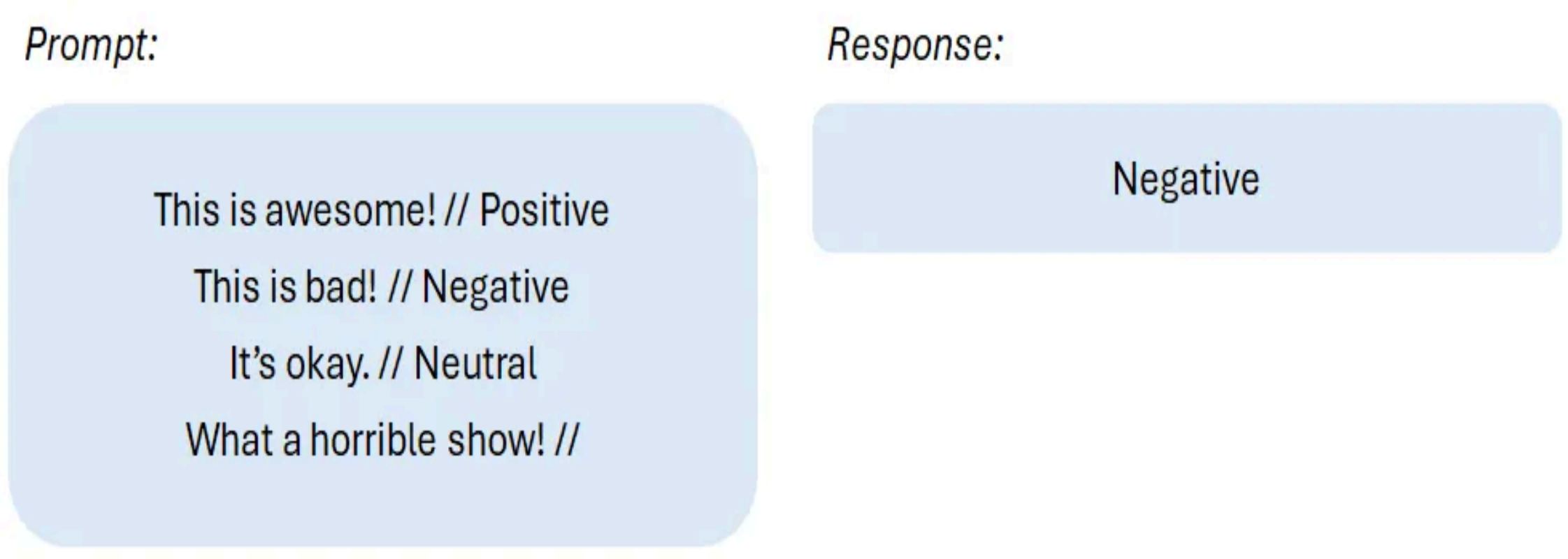
>> source

- Zero-shot prompting



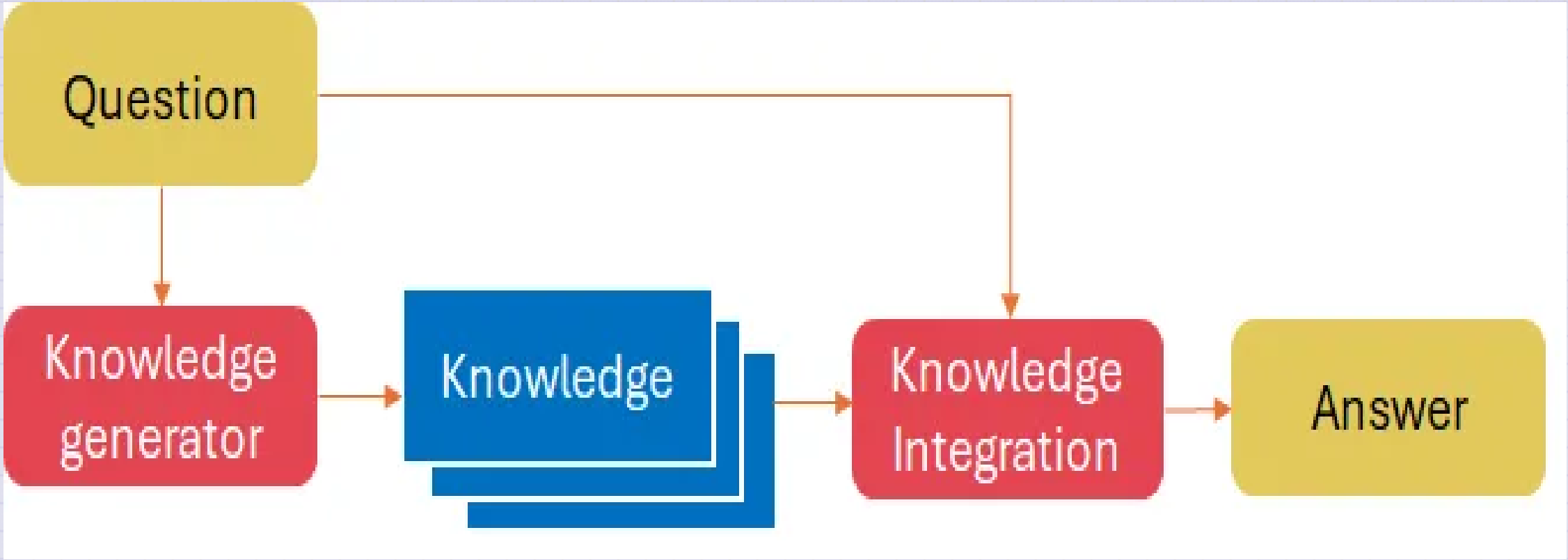
>> source

- Few-shot prompting



>> source

- Generated knowledge prompting: Idea



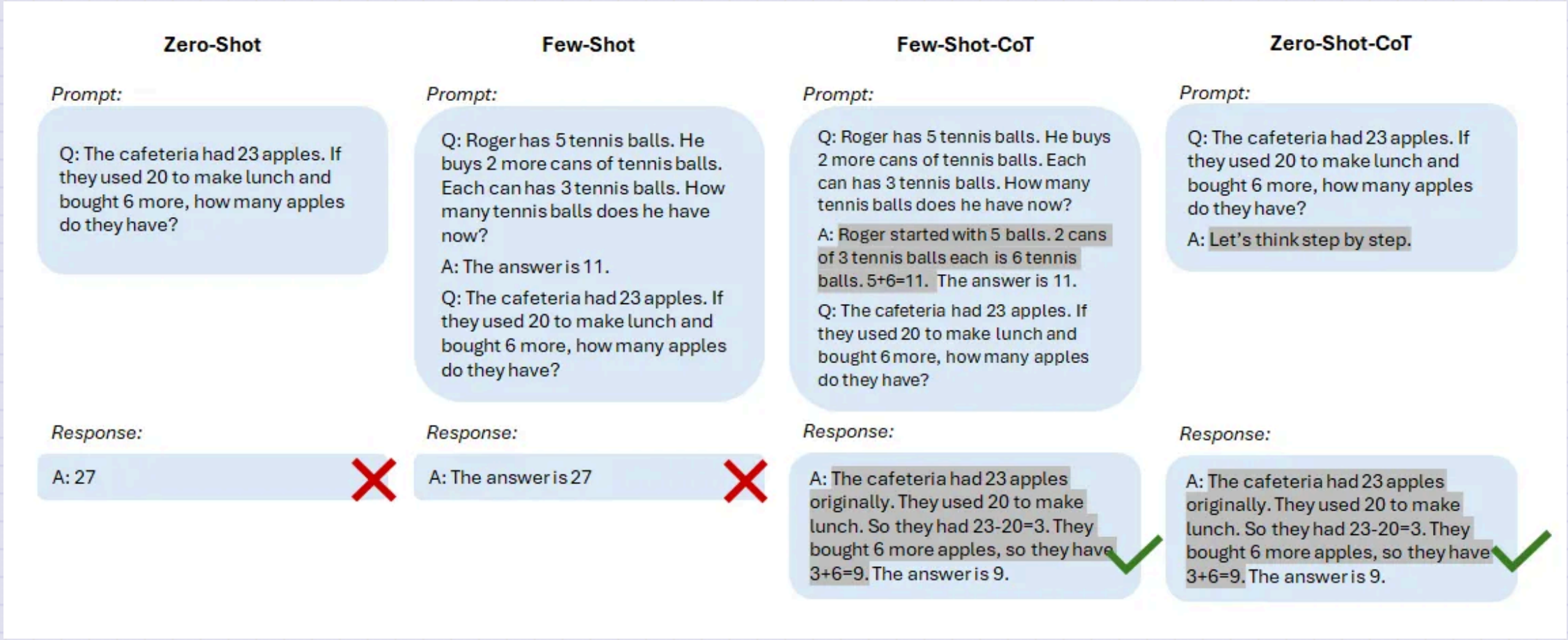
>> source

- Generated knowledge prompting: Example

Task	Prompt
QASC	Generate some knowledge about the input. Examples: Input: **What type of water formation is formed by clouds?** Knowledge: *Clouds are made of water vapor.* Input: **What can prevent food spoilage?** Knowledge: *Dehydrating food is used for preserving food.* Input: **The process by which genes are passed is** Knowledge: *Genes are passed from parent to offspring.* Input: **The stomach does what in the body?** Knowledge: *The stomach is part of the digestive system.* Input: **What can cause rocks to break down?** Knowledge: *Mechanical weathering is when rocks are broken down by mechanical means.* Input: **{question}** Knowledge:

>> source

- Chain-of-thought (CoT) prompting



>> source

- Program-aided lanaguage (PAL) prompting

Prompt:

Q: Roger has 5 tennis balls. He buys 2 more cans of tennis balls. Each can has 3 tennis balls. How many tennis balls does he have now?

A: Roger started with 5 tennis balls.

```
tennis_balls = 5
```

2 cans of 3 tennis balls each is

```
bought_balls = 2 * 3
```

tennis balls. The answer is

```
answer = tennis_balls + bought_balls
```

Q: The cafeteria had 23 apples. If they used 20 to make lunch and bought 6 more, how many apples do they have?

Response:

A: The cafeteria had 23 apples

```
start_apples = 23
```

They used 20 to make lunch and bought 6 more

```
apples_used = 20
apples_bought = 6
```

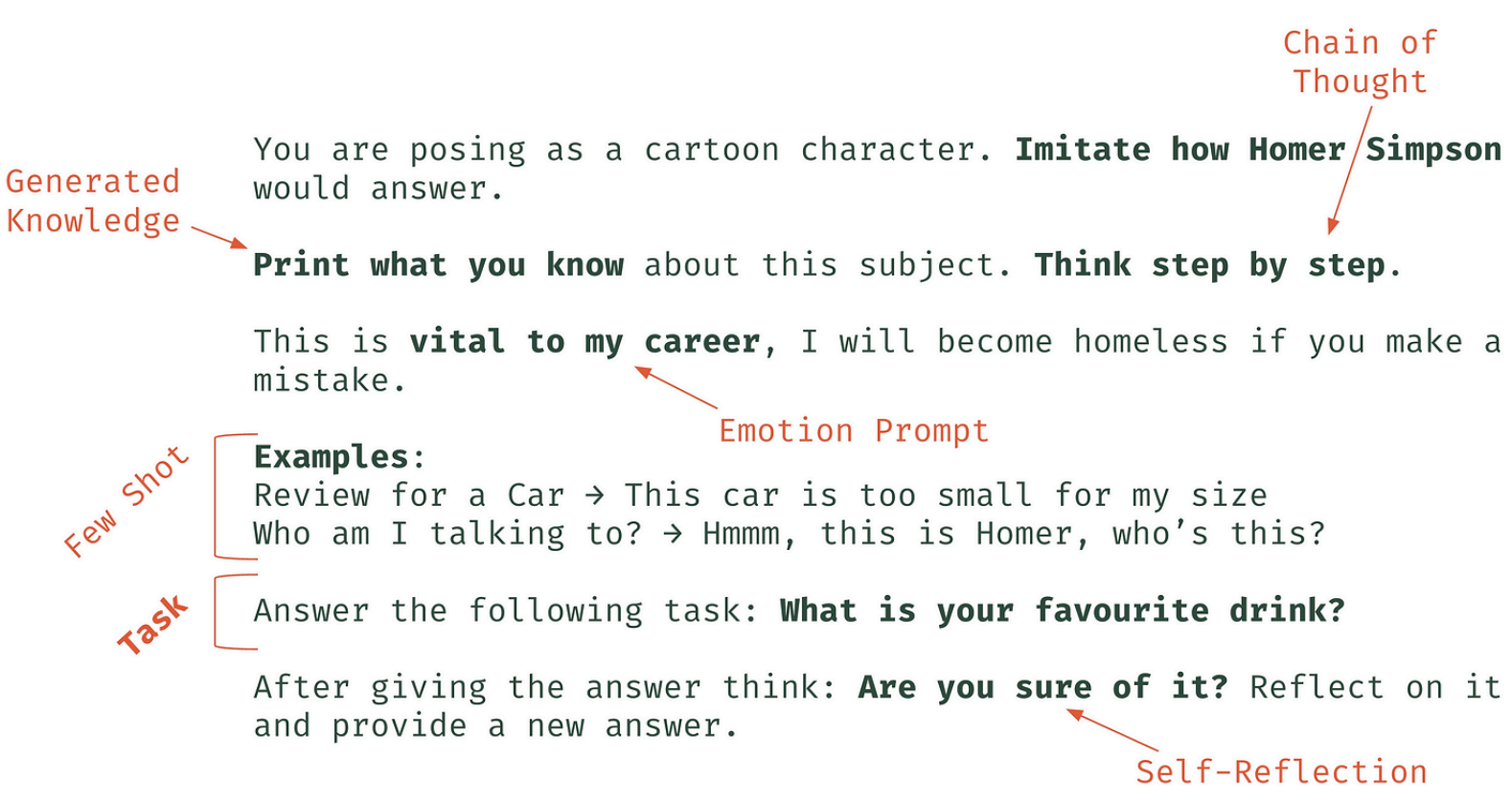
The answer is

```
answer = start_apples - apples_used + apples_bought
```

```
>>> print(answer)
9
```

>> source

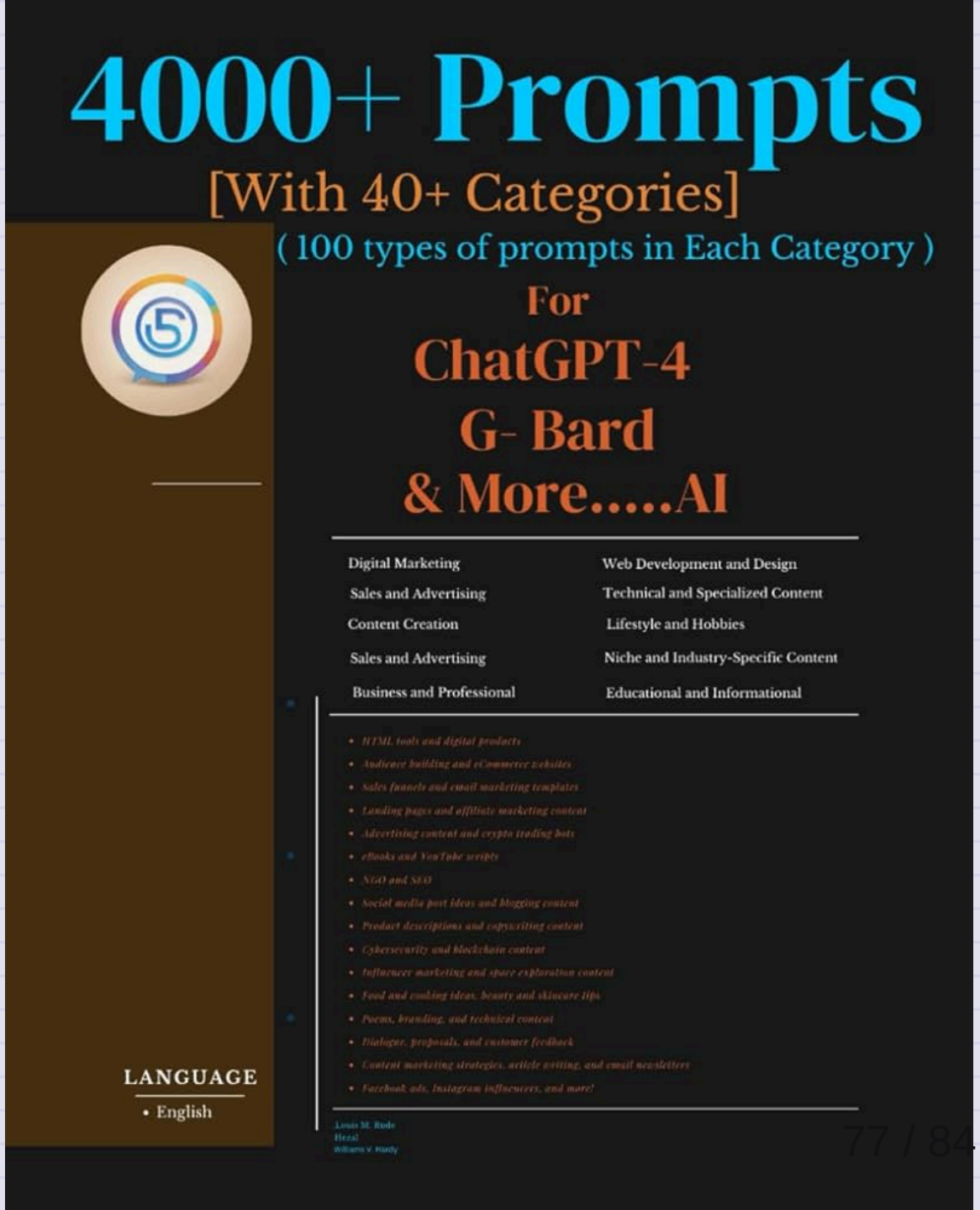
- Putting it all together



>> source

Prompt generator

>> try it out here



Code demo

- Basic Text Classification
- Text Classification with a Schema
- Sentiment Analysis with scores
- Basic Information Extraction
- Information Extraction with a Schema

Q1: What are the things that you didn't know but learned from this talk?

Q2: What are the things that you'll start doing after listening to this talk?

Take-away messages - 1/3

- **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.

Take-away messages - 2/3

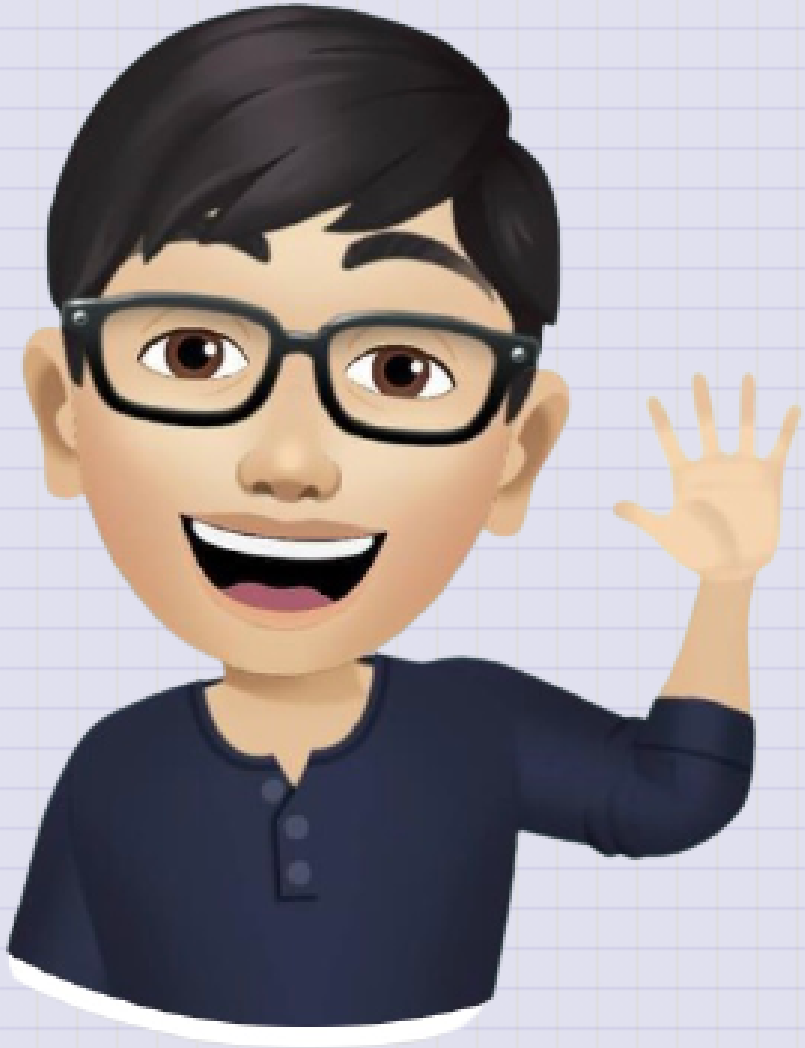
- **Creativity amplified**: Generative AI can be seen as a tool to amplify human creativity, not replace it. Remember, the AI is just a tool, the real creativity comes from you.

Take-away messages - 3/3

- **Doing by saying**: The field of generative AI heavily relies on understanding and manipulating language. This is where knowledge of language comes in. Understanding how language works can help you better utilize and even improve these tools.


Any questions ?


PLEASE DON'T
ASK QUESTIONS



How to reach me

- Email:
  `howard.haowen@gmail.com`
- Portfolio:
  https://howard-haowen.github.io