

# Prompt Engineering the Right Way

Kellogg Research Support

---

**October 18, 2024**

Northwestern | Kellogg

# Roadmap

## Advice

- ✓ Prompt Terminology
- ✓ Some Tips and Tricks
- ✓ Testing in the **Playground**

## Hands-On Projects

- ✓ Female Participation in BoD Meetings
- ✓ Classifying Protest Activity in Front Page Images
- ✓ Cybersecurity in 10K Risk Statements

## Scaling Up

- ✓ **Marketing Study with GPT Personas**
- ✓ Using the Batch API

All materials provided here: [https://github.com/rs-kellogg/prompting\\_right\\_way/](https://github.com/rs-kellogg/prompting_right_way/)

# The Art of Prompting

**Prompting** – giving an LLM an instruction or question to elicit a response

**Prompt Crafting** – tailoring a prompt in real-time to generate a relevant response for a **one-off, specific instance**

**Prompt Engineering** - curating successful prompts as scalable solutions for **diverse use cases possibly with future unknown inputs**



Definitions from: Fortune Magazine Presentation <https://www.youtube.com/watch?v=AxfmzLz9xXM>

# Why is it Useful?

## To Prevent Problematic LLM Responses:

- ❖ Too Vague or Too Broad
- ❖ Hallucinations / Wrong Answers
- ❖ Biased
- ❖ Unusable Format

# PROMPT TERMINOLOGY

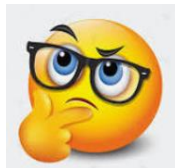
# Zero-Shot vs Few-Shot Prompting

**Zero-Shot** - Directly asking the model to perform a task without providing examples.

**Few-Shot** - Offering examples within the prompt to guide the model toward the desired output.

## Few Shot Uses:

1. Consistently Formatted Output\*
2. Clarify Ambiguous Task



*Why provide examples if Zero-Shot prompts work?*

Exactly! Pricing is based on token inputs, so a few-shot prompt is more costly. Avoid it when a zero-shot prompt produces usable responses.

\* Note – Possible to set response formatting in the API, including json schema.

# Zero-Shot Example

## User Prompt:

Rate this vehicle on a scale of 1 to 10 on appearance, durability, and affordability based on this review:



Image from:

<https://www.caranddriver.com/tesla/cybertruck>

*The Tesla Cybertruck's bold, futuristic appearance is polarizing, with its angular, stainless-steel design resembling something out of a sci-fi movie. While its starting price is competitive for an electric truck, higher-end models with full features can get pricey. It excels in functionality with impressive towing capacity, off-road capabilities, and cutting-edge tech, making it both rugged and innovative.*

## Response:

Based on the review provided, here is a rating for the Tesla Cybertruck on a scale of 1 to 10....

1. **\*\*Appearance\*\***: 8/10

- The Cybertruck's bold and futuristic design.....

2. **\*\*Durability\*\***: 9/10

- The use of stainless steel and the mention of its rugged capabilities....

3. **\*\*Affordability\*\***: 7/10

The starting price is noted as competitive for an electric truck, which.....

# Few-Shot Example

## User Prompt:

Here's an example of how to rate a vehicle review on a scale of 1 to 10 on appearance, durability, and affordability:

*Review: 'The Honda Accord's sleek design is appealing but not groundbreaking. It's known for its durability and long lifespan, though its higher-end trims can be expensive.'*

Rating: Appearance: 7/10; Durability: 9/10; Affordability: 6/10



Rate this review using the same format:

*Review: The Tesla Cybertruck's bold, futuristic appearance is polarizing, with its angular, stainless-steel design resembling something out of a sci-fi movie. While its starting price is competitive for an electric truck, higher-end models with full features can get pricey. It excels in functionality with impressive towing capacity, off-road capabilities, and cutting-edge tech, making it both rugged and innovative.*

## Response:

Rating: Appearance: 8/10; Durability: 8/10; Affordability: 5/10

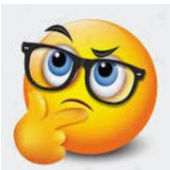


# System vs User Prompts

A **User Prompt** typically consist of an input or query. While the **System Prompt** defines the **role**, provides **context**, or sets the **tone** for the model.

## System Prompt Uses:

1. Role Assignment – the persona to adopt
2. Context Setting – specifies subject matter or topic
3. Tone/Style Guidance, e.g., be concise



*Can't I place the System Prompt in the User Prompt?*

Developers give us insight on how these models are trained. You might get the same results if you diverge from their guidance, or you might not.

# System Prompt Example

**User Prompt:** What is the best way for the government to reduce unemployment during a recession?

## Response:

1. ***Increase Public Spending:*** ...
2. ***Extend Unemployment Benefits:*** ....
3. ***Tax Cuts and Incentives:*** ....
4. ***Support Small and Medium Enterprises (SMEs):*** ...
5. ***Workforce Training and Education:*** ....
6. ***Monetary Policy Measures:*** ...
7. ***Boosting Consumer Confidence*** ...

# System Prompt Example

**User Prompt:** What is the best way for the government to reduce unemployment during a recession?

**System Prompt:** You are a Chicago School economist who strictly follows Milton Friedman's principles. You believe recessions are caused by poor monetary policy, not by a lack of government spending.

## Response:

*As someone who adheres to Milton Friedman's principles, I believe that the key to reducing unemployment during a recession is through **sound monetary policy rather than increased government spending**. The government should focus on ensuring that the money supply grows at a stable and predictable rate.....*

# Prompt Chaining

Involves breaking a complex task into several linked prompts where the output from one prompt becomes the input for the next.

## Especially Useful if you:

1. introduce new inputs after receiving the output of a step – combine with other sources or code
2. perform subsequent steps **conditional** on a previous response

# Prompt Chaining Example

**User Prompt 1:** Is this image a deed? Answer Yes or No only.

**Response 1:** Yes

If No:

Stop

If Yes:

OCR Image through Textract

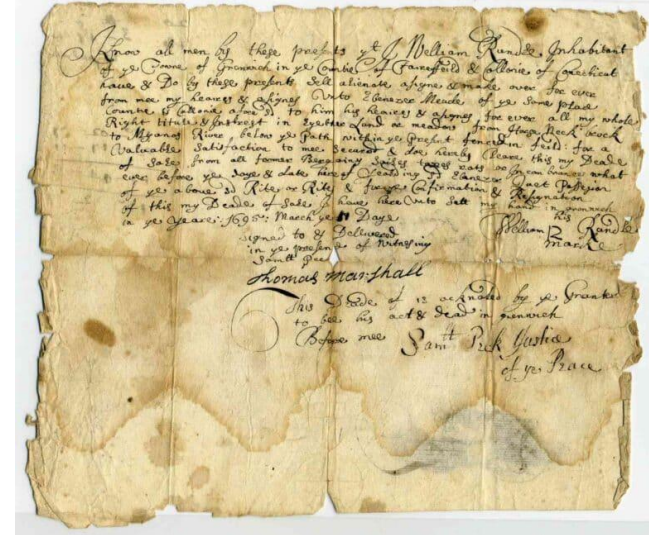


Image from: <https://greenwichhistory.org/how-to-read-an-early-american-real-property-deed/>

**User Prompt 2:** Here's the text for historic deed:

"Know all men by these presents & I William Rundle Inhabitant of ye Towne of Greenwich in ye County of Fairfield & Collonie of Connecticut....."

Determine the owner, property, city, state, value, date....

# Chain of Thought Prompting

Ask the model to think through its reasoning “**step-by-step**”

--- OR ---

Explicitly provide it with steps

**Especially Useful for:**

- 1. Complex Problem Solving** - Breaking down multifaceted tasks.
- 2. Performing Calculations** - Ensuring accurate numeric assessments.
- 3. Understanding Model Handling of Edge Cases** - Clarifying responses in atypical scenarios.

# Chain of Thought Example

## Task – Estimate Private Jet Expenses for Execs

Name and Principal Position	Year	Salary (\$)	Bonus (\$)	Stock Awards (\$)	All Other Compensation (\$)	Total (\$)
Marc Rowan, Chief Executive Officer	2023	100,000	—	—	220,760	320,760
	2022	100,000	—	—	210,011	310,011
	2021	100,000	—	—	202,310	302,310
Martin Kelly, Chief Financial Officer	2023	1,000,000	—	1,043,706	1,626,223	3,669,929
	2022	1,000,000	—	1,082,082	1,491,637	3,573,719
	2021	1,000,000	—	19,160,680	4,135,828	24,296,508
Scott Kleinman, Co-President of AAM	2023	100,000	—	4,219,954	11,075,543	15,395,497
	2022	100,000	—	3,976,560	8,081,234	12,157,794
	2021	1,200,000	—	392,478,845	43,308,266	436,987,111
James Belardi, Chairman, Chief Executive Officer and Chief Investment Officer of AHL	2023	1,875,000	— <sup>(3)</sup>	6,690,787	42,216,981	50,782,768
	2022	1,741,141	209,326	6,428,865	36,316,113	44,695,445
John Suydam, Chief Legal Officer of AGM (through December 31, 2023)	2023	2,500,000	—	4,524,885	2,718,116	9,743,001
	2022	2,500,000	—	3,961,338	1,928,714	8,390,052

- (1) Represents the aggregate grant date fair value of stock awards granted, as applicable, computed in accordance with FASB ASC Topic 718. For Mr. Belardi, includes both the annual incentive bonus RSUs and the long-term incentive RSUs granted to him in 2023. For Mr. Kleinman, amounts shown relate to legacy performance fee programs in which he had vested before 2023 that require that a portion of the performance fee amounts be used to purchase restricted shares of common stock. The amounts shown in this column do not reflect compensation actually received by the named executive officers, but instead represent the aggregate grant date fair value of the awards. See note 16 to our consolidated financial statements in our Annual Report on Form 10-K for the year ended December 31, 2023 for further information concerning the assumptions made in valuing our RSU awards.
- (2) Amounts included for 2023 represent, in part, actual incentive pool cash distributions of \$1,175,000 for Mr. Kelly, realized carry cash distributions of \$199,365 for Mr. Kelly, \$11,067,041 for Mr. Kleinman and \$2,465,785 for Mr. Suydam and partner benefits stipends of \$250,000 for each of Messrs. Kelly and Suydam. For Mr. Belardi, the amount includes distributions on his ISG partnership interest totaling \$40,989,467, amounts in respect of his ISGI profits entitlement (\$1,184,117), and fees paid by AHL for tax preparation services (\$23,597). AHL maintains a corporate aircraft for efficiency and business planning purposes. Mr. Belardi used the corporate aircraft for two **personal flights** in 2023 and fully reimbursed AHL for this personal use. Accordingly, no amount is reflected for such use. Personal use of the AHL corporate aircraft is subject to a formal policy that was approved by the Compensation Committee that sets forth the criteria and procedures applicable to its use. Mr. Belardi and AHL have entered into a time-sharing agreement, pursuant to which Mr. Belardi may use the corporate aircraft for up to 25 flight hours per year, provided that the number of flight hours and other incidentals under such agreement shall be further limited so that the amount of payments from Mr. Belardi pursuant to such agreement (including such tax payments)

**Apollo Global Management Proxy Statement from:**

<https://ir.apollo.com/sec-filings/content/0001193125-24-118090/d802194ddef14a.htm>

# Chain of Thought Example

## User Prompt Example:

Using the executive compensation table, estimate the corporate jet costs for any executive by following these steps:

- 1. Determine Usage:** Identify the number of personal flights taken by the executive.
- 2. Reimbursement Check:** Note if the executive reimbursed the company for personal use.
- 3. Estimate Costs:**
  - Fuel:** \$1,500 per flight
  - Maintenance & Crew (M&C):** \$2,000 per flight
  - Overhead:** \$500 per flight
- 4. Calculate Total Cost:**

$$\text{Total Cost} = \# \text{ of Flights} * (\text{Fuel} + M \& C + \text{Overhead})$$



# Tips and Tricks

- Be clear, detailed but direct, and unambiguous
- Don't ask an LLM to do something you can program
- Give the LLM a small, digestible input
- Take advantage of **System Prompts**
- Provide examples of desired responses: **few shot**
- Ask the LLM to work **step-by-step** or provide steps.
- Build on API responses with **prompt chaining**
- **Simplicity and repetition are key!**
- Reformulate prompts as much as needed in **Playground!**

# THE PLAYGROUND

# Chat

What can I help with?



Message ChatGPT



Create image



Get advice



Code



Brainstorm

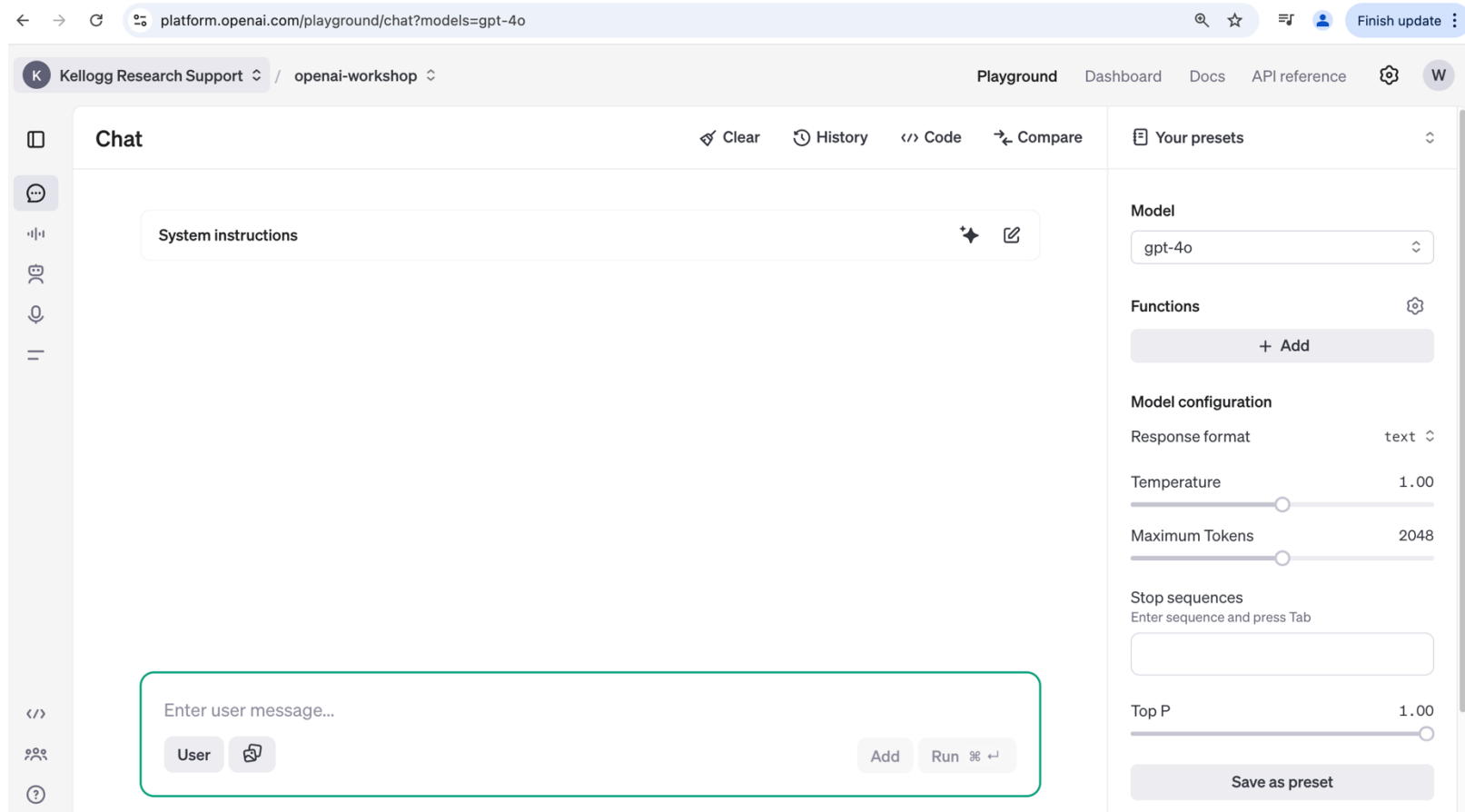


Summarize text

More

Best for quick, one-off, conversational interactions

# Playground



Ideal for detailed customization, testing, and experimenting with prompts.

# Playground Tour

<https://platform.openai.com/playground/>

The screenshot shows the OpenAI Playground interface with the following numbered annotations:

- (1) Left sidebar menu containing icons for Chat, System instructions, History, Code, Compare, and Presets.
- (2) System instructions text area.
- (3) User message input area with a text box and a "User" button.
- (4) Model selection dropdown menu.
- (5) Model configuration panel including Response format, Temperature, Maximum Tokens, Stop sequences, and Top P.
- (6) History button.
- (7) Code button.
- (8) Run button.

# (1) Interface Modes

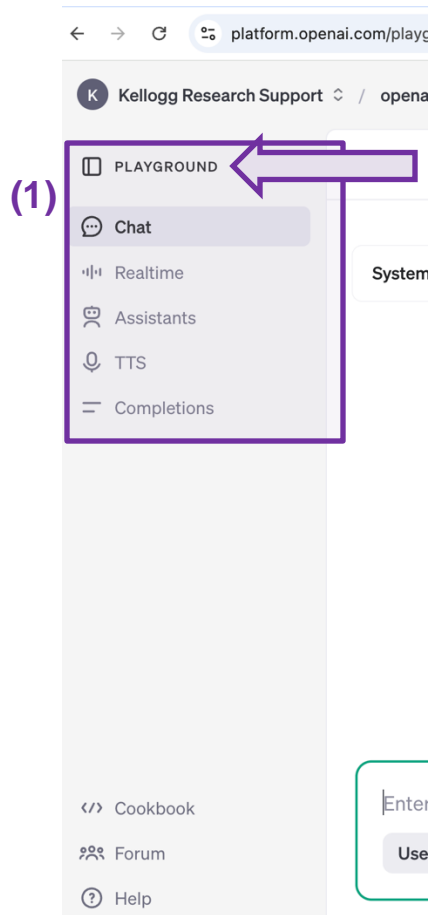
**Chat** - Interactive mode for back-and-forth conversations with the model.

**Realtime** - Generates responses instantly as you type for dynamic exchanges.

**Assistants** - Pre-configured for specific tasks

**TTS (Text-to-Speech)** - Converts text inputs into spoken language.

**Completions** - Generates text to complete or expand on a given prompt.



## (2) System and (3) User Prompts

The screenshot displays the OpenAI Playground interface at `platform.openai.com/playground/chat?models=gpt-4o`. The interface is divided into three main sections: a left sidebar, a central chat area, and a right sidebar for model configuration.

**Central Chat Area:**

- (2) System instructions:** A text box at the top of the chat area, highlighted with a purple border, containing the text "System instructions".
- (3) User message:** A text box at the bottom of the chat area, highlighted with a purple border, containing the placeholder text "Enter user message...". Below this box are buttons for "User", "Add", and "Run".

**Right Sidebar (Model Configuration):**

- Model:** A dropdown menu set to "gpt-4o".
- Functions:** A button labeled "+ Add".
- Model configuration:**
  - Response format:** A dropdown menu set to "text".
  - Temperature:** A slider set to 1.00.
  - Maximum Tokens:** A slider set to 2048.
  - Stop sequences:** A text box with the placeholder "Enter sequence and press Tab".
  - Top P:** A slider set to 1.00.
  - Save as preset:** A button at the bottom.

**Top Navigation Bar:**

- Buttons for "Clear", "History", "Code", and "Compare".
- Buttons for "Your presets", "Dashboard", "Docs", "API reference", and "W".
- A "Finish update" button.

# (4) Model Selection and (5) Presets

## (4) Model Selection –

Select any model from the dropdown.

**We use gpt-4o for all examples.**

## (5) Presets

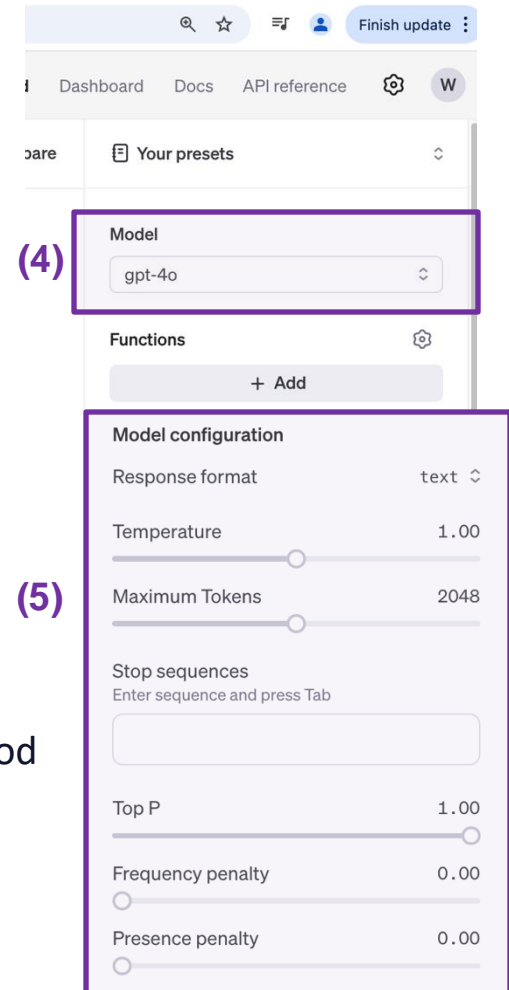
**Temperature** - Controls the randomness of responses.  
(Set to 0 for same response each time)

**Max Tokens** - Sets the maximum length for the model's output.

**Top P (Nucleus Sampling)** - Dictates the variety in responses by only considering the top 'P' percent of probable words.

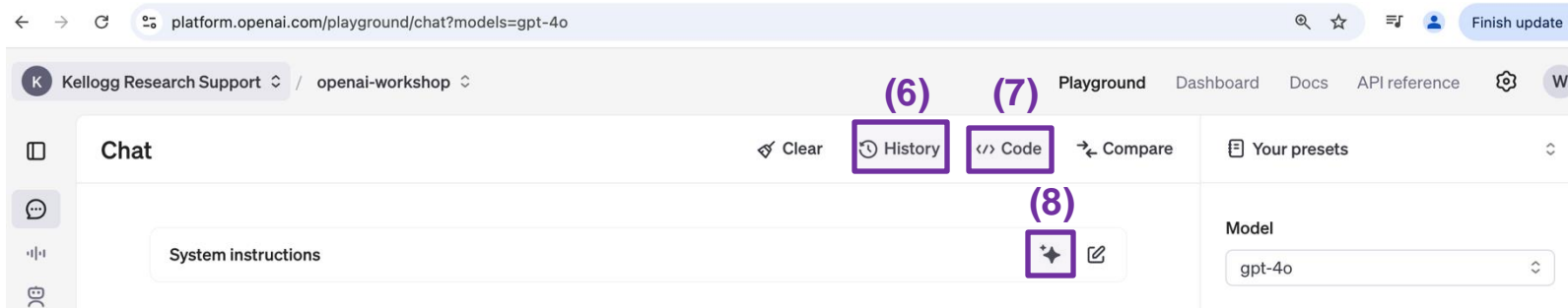
**Frequency Penalty** - Reduces repetition by decreasing the likelihood of frequently used words.

**Presence Penalty** - Promotes the introduction of new topics in the conversation.





# (6) History, (7) Code, (8) Prompt Generator



**(6) History** – Shows the history of all prompt/responses made. *Note that prompts with attached images will not be saved.*

**(7) Code** – Provides a code snippet to get started running your prompt through the API

**(8) Prompt Generator** – Input what you'd like to do, and the generator will translate this to a prompt you can use.\*

\* Beyond Prompt Generators, **DSPy** is a tool that simplifies building machine learning models by letting you describe your needs in plain language, while generating internal prompts not visible to users described here: <https://github.com/stanfordnlp/dspy>.

# HANDS-ON PROJECTS

# Ex 1 – Female Participation in Boards

Do women participate the same way as men?

**Allan Edwards (Chairperson):**

"Let's move on to the updates from our departments. I think it's important to stay aligned on our ongoing projects. John, can you start with the finance update?"

**James Rodriguez (CFO):**

"Sure, Alan. We've finalized the budget allocations for next quarter, and everything is looking stable. We're slightly under on office supply expenses this month, but nothing significant."

**Allan Edwards:**

"Good to hear. Any concerns?"

**Joe Doe (CEO):**

"None at the moment."

**Jane Smith (Marketing Director):** .....

*Find text in the the **news\_ex** folder of the git repo.*

# Ex 2 – Violence in Front Page Images

Is this a peaceful or violent event based on the image?



Find images in the the **news\_ex** folder of the git repo.

Front Pages for Jan 6 taken from: <https://worldcrunch.com/world-affairs/capitol-insurrection-26-front-pages-from-around-the-world>

Front Pages for Ferguson taken from: <https://www.poynter.org/reporting-editing/2020/in-nearly-every-state-front-pages-capture-outrage-after-george-floyds-death/>

# Ex 3 – Score Cybersecurity Risk in 10Ks

Provide a rating for Cybersecurity Risk:

*In evaluating our business landscape, we recognize several potential threats that could impact our operations and market position. Increasing competition from both established players and emerging startups, particularly in the technology sector, could erode our market share and pressure profit margins. Additionally, shifts in consumer preferences towards more sustainable products may challenge our traditional offerings, requiring adaptation to meet evolving demands. We also face potential regulatory changes that could impose stricter compliance requirements, along with the need to address cybersecurity risks in an increasingly digital environment.*

.....

Find text in the the **risk\_ex** folder of the git repo.

Similar concept to this study: <https://pubsonline.informs.org/doi/full/10.1287/mnsc.2022.02056>

# SCALING UP – BATCH API



# Marketing Study with GPT Personas

We asked **DALL-E** to produce 3 soft drink ads for us:

1. with cute kawaii characters
2. with people at a Cubs game enjoying the drink
3. from a hopeless dystopian future, where the soft drink looks like it might taste good



Find images in the the **mkty\_ex** folder of the git repo.

# Marketing Study with GPT Personas

We then asked GPT 4o to create 50 personas based on demographic information we gave it on race and education levels in Monkey's Elbow, Kentucky:

It produced personas like this:

*David Smith is a 25-year-old actuary who relaxes by watching Anime while knitting.*

To scale up, we provide it with a persona and ask:

*Which image do you prefer and why?*



# SUMMARY

# Take-Aways

## ✓ Do

Be detailed, clear, and unambiguous

Take advantage of the System Prompt

Provide Examples

Give instructions step-by-step

Experiment with prompts in the Playground!

## X Don't

Ask the model to do things you can program yourself

Give the model too large of an input or too many tasks at once

# Reference Materials

## **Open AI Prompt Engineering Guide:**

<https://github.com/nuitrcs/promptEngineering>

## **NUIT Prompt Engineering Workshop:**

<https://github.com/nuitrcs/promptEngineering>

## **NUIT Prompt Engineering Cheat Sheet:**

[https://github.com/nuitrcs/promptEngineering/blob/main/prompt\\_engineering\\_cheat\\_sheet.pdf](https://github.com/nuitrcs/promptEngineering/blob/main/prompt_engineering_cheat_sheet.pdf)

# Appendix - Advanced Prompting

## **Retrieval Augmented Generation:**

<https://www.promptingguide.ai/techniques/rag>

*(now available in the Assistant API Playground)*

## **ReAct Prompting:**

<https://www.promptingguide.ai/techniques/react>

## **Directional Stimulus Prompting:**

<https://www.promptingguide.ai/techniques/dsp>