



# Unlocking Visual Intelligence: Advanced Prompt Engineering for Vision-Language Models

Alina Li Zhang

Senior Data Scientist

LinkedIn Learning

# Advanced Prompt Engineering for Vision-Language Models



What will happen in the next second?

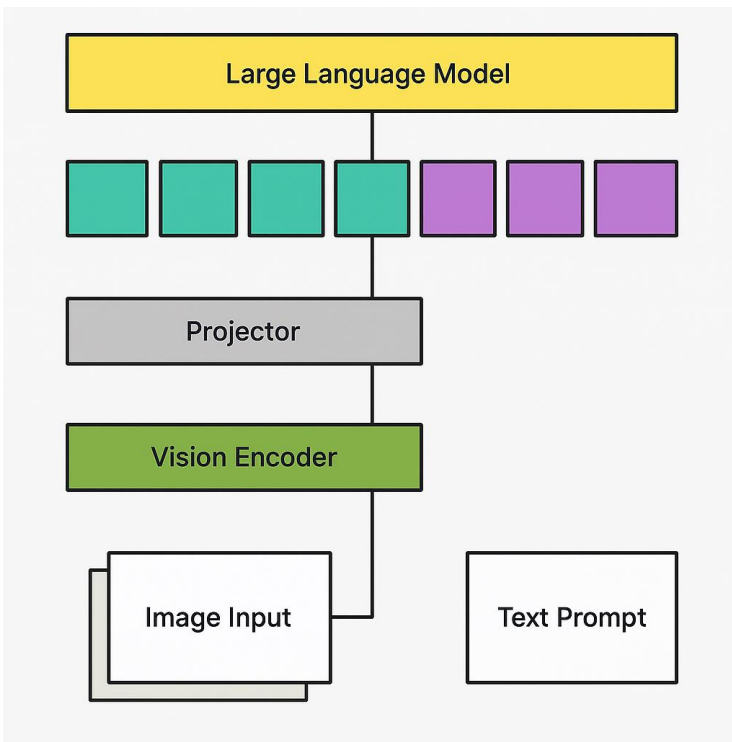
# Beyond Text: The Era of AI with Eyes



GPT-4o:

“The cat is likely to knock over the glass of water. Cats are known for their playful and mischievous behavior, and this one appears to be testing the glass with its paw.”

# What are Vision-Language Models (VLMs)?



“Sight turned into insight;  
Seeing became understanding;  
Understanding led to action.”

– Fei Fei Li

# Talk to AI: What is Prompt Engineering?

**Prompt engineering** is the art and science of designing effective prompts to guide AI models toward generating accurate, relevant, and desired outputs.

## Why prompt engineering matters:

- Improving output quality and relevance
- Enhancing control and predictability
- Enabling flexibility and reusability



# The Pentagram Framework for Prompt Engineering



# Prompt Engineering: Persona



## Persona:

You are a seasoned professor at Stanford University, specializing in business management with extensive experience in teaching MBA students.

# Prompt Engineering: Context



**Context:**  
The blog is intended for your MBA students, who are well-versed in business concepts but may lack practical experience.

# Prompt Engineering: Task



## Task:

Write a blog post to guide students on how to formulate effective action items that are clear, achievable, and measurable.

# Prompt Engineering: Output



**Output:**  
The tone should be professional but engaging, encouraging students to apply these concepts in their project.

# Prompt Engineering: Constraint



**Constraint:**  
Avoid overly technical jargon that might confuse MBA students who are not yet industry professionals.

# The Pentagram Framework for Shop Safe AI

How to instruct AI models to detect suspicious behaviors in a grocery store?





## Persona

You are an AI security analyst specializing in monitoring grocery store surveillance footage. Your expertise lies in detecting suspicious behaviors such as theft, item concealment, and price tag tampering.



## Context

You operate within a real-time video surveillance system for retail environments.

Users will provide one or more extracted frames from security camera footage for analysis.

The store layouts typically include aisles, checkout zones, and product shelves.



## Task

Analyze the provided frame(s) to identify any suspicious or unlawful behavior.

For each detected incident, determine the subject involved, describe their actions, assess the risk level, and recommend appropriate responses.



## Output

Present your findings in a structured, dashboard-style alert format using clear and concise language. Incorporate relevant emojis to enhance readability.

## Few-shot learning



For example:

### Theft Alert – Self-Checkout Zone

- Subject: Male, dark hoodie, jeans
- Item: Yellow packaged goods
- Behavior: Running past self-checkout area with item, no scan recorded
- Confidence: 98.9%
- Risk Level: HIGH
- Action: Initiate immediate floor alert. Block exit routes if safe. Notify security or law enforcement



## Constraint

- Rely only on visible evidence and avoid assumptions.
- Keep reports clear and concise.
- Follow privacy and security policies.
- Focus on actionable insights.

# Testing: Concealment Alert



## ⚠ Concealment Alert

**Subject:** Elderly woman,  
beige jacket, gray handbag

**Item:** Small jar

**Behavior:** Placing item  
directly into personal  
handbag

**Confidence:** 92.4%

**Risk Level:** MEDIUM

**Action:** Monitor subject at  
checkout. If item is not  
scanned, flag for follow-up

# Testing: Price Tag Tampering



## ⚠️ Price Tag Tampering

Subject: Male, denim shirt

Item: Jar of pasta sauce

Behavior: Swapping barcode/  
price label on the product

Confidence: 95.1%

Risk Level: MEDIUM

Action: Observe checkout  
behavior for price mismatch.  
Flag product for audit

# AI Eyes on the World: More Applications



“My sword I leave to him who can wield it.”

– Charlie Munger



- What are VLMs?
- What is the Pentagram Framework for prompt engineering?
- Practical example: How to craft the Pentagram prompt for a Shop Safe AI system?

# What's Next?



Course **Popular**

## Build Your Own GPTs

46m • Beginner • Feb 2024

✓ Completed 12/2/2024



LinkedIn • By: Alina Zhang

4.7 ★★★★★ (2,188) • 115,342 learners • Certificate Eligible

