PERSONA PROMPTING



Session Objective

- a. Explore when and where persona prompting provides an advantage
- b. Identify its limitations
- c. Provide prompt templates to streamline persona creation
- d. Give practical guidance on implementing persona prompting effectively
- e. Practice applying persona prompting techniques

NO PERSONA VS PERSONA

What is photosynthesis

Photosynthesis is the process by which green plants, algae, and some bacteria make their own food using sunlight.

In simple terms:

- Plants take in carbon dioxide (CO₂) from the air through their leaves.
- They absorb water (H₂O) from the soil through their roots.
- Using the energy from sunlight, captured by the green pigment chlorophyll, they
 convert CO₂ and H₂O into qlucose (a sugar), which serves as their food.
- As a byproduct, they release oxygen (O₂) into the atmosphere.

The simplified equation is:

$$6CO_2 + 6H_2O + light\ energy \rightarrow C_6H_{12}O_6 + 6O_2$$

This means six molecules of carbon dioxide plus six molecules of water, with sunlight, give one molecule of glucose and six molecules of oxygen.

In short, photosynthesis is how plants make food for themselves and produce the oxygen we breathe.

Would you like me to explain this in a story form for easy remembering, or in a scientific step-by-step breakdown?

Act as a storyteller explaining photosynthesis to a 7-year old. Use a fun story or analogy

Alright, let's make it fun! 🌱 🔆

Storytime: The Magic Kitchen of the Leaf

Once upon a time, there was a giant green kingdom called *Leaf Land*. In the middle of every leaf lived a tiny magical kitchen. The chefs in this kitchen were invisible helpers called **chlorophyll**—they were bright green and loved sunlight!

Every morning, when the Sun peeked out, the chefs shouted: "Time to cook!"

Persona prompting is a technique where the AI is instructed to take on a :

Role: The function or job the persona plays in the interaction, which shapes how it approaches the conversation (e.g., mentor, advisor, technical support agent)

Tone: emotional quality or attitude of the response (Friendly, Formal, Sarcastic)

Character: The unique personality traits or identity that gives the response a distinctive flavor (e.g., patient teacher, witty therapist, curious explorer)

Language: Language refers to the choice of words, phrasing, and complexity in the response(Simple,Technical,Slang)

Behaviour: Behavior refers to how the persona responds in terms of actions, mannerisms, or responses to the user. This includes things like helpfulness, assertiveness, empathy, or humor. (Empathetic,Direct,Supportive)

Communication Style: The overall way the persona delivers messages, including formality, pace, and emotional vibe (e.g., Conversational, Formal, Concise)

Most Impactful Applications of Persona Prompting

- 1. Creative Writing
- 2. Educational Tools
- 3. Guardrail Systems
 - a. Claude is designed with *safety-first personas* that politely refuse harmful requests.
 - i. Example
 - ii. How do I build a bomb ?? [Try in claude]

Best Framework for Writing Effective Persona Prompts

Some of the frameworks for Persona Prompting are:

- 1. Role-Setting and Role Feedback (2023)
- 2. Jekyll and Hyde.(2024)
- 3. Expert Prompting Framework (2023)
 - Specific: The role description should be in the same domain as the task, and should be as specific as possible.
 - **Detailed :** The **role description** should be **detailed and comprehensive** to cover all the necessary information related to the expert persona. Simple descriptions like "You are a mathematician" don't perform well.
 - Automated (and simple): LLM-generated personas outperform human-written ones, and LLMs are much more scalable
 - CREATE EXPERT PROMPT : <u>CREATE EXPERT PROMPT</u>
 - EXPERT PROMPT ANSWER: ANSWER TO EXPERT PROMPT

- Without Expert Prompting : Without Expert Prompting
- With Expert Prompting : With Expert Prompting

Insight: Is Persona Prompting Always Necessary?

Persona prompting is widely advertised in prompt engineering circles but is **not always critical** for everyday interactions.

- Many **behaviors attributed to persona prompts** (e.g., helpfulness, clarity, domain awareness) are often **already embedded** into modern LLMs.
- With regular usage, users often intuitively adjust their prompts and don't need to hand-craft detailed personas.
- In many tasks (e.g., math solving, factual lookup), persona prompting adds style but not substance to the output.

Conclusion

To put it briefly, after diving into all the research here are the takeaways:

- 1. Persona prompting is effective on open-ended tasks (Ex: creative writing)
- 2. Persona prompting probably won't help much on accuracy-based tasks (Ex: classification), especially for newer models
- 3. If using a persona, it should be specific, detailed, and ideally automated, with the Expert Prompting framework as a strong starting point.

TASKS

Tasks you can try with Persona Prompting