Frameworks for Prompt Writing and Model Control in Generative AI

WRITTEN BY: UTKARSH KUMAR DUBEY

WRITTEN AS A RESEARCH SUPPORT BACKUP FOR FOSSEE

Frameworks for Prompt Writing and Model Control in Generative AI

written by

Utkarsh Kumar Dubey

Final Year CS Engineering Student

VIT Bhopal University

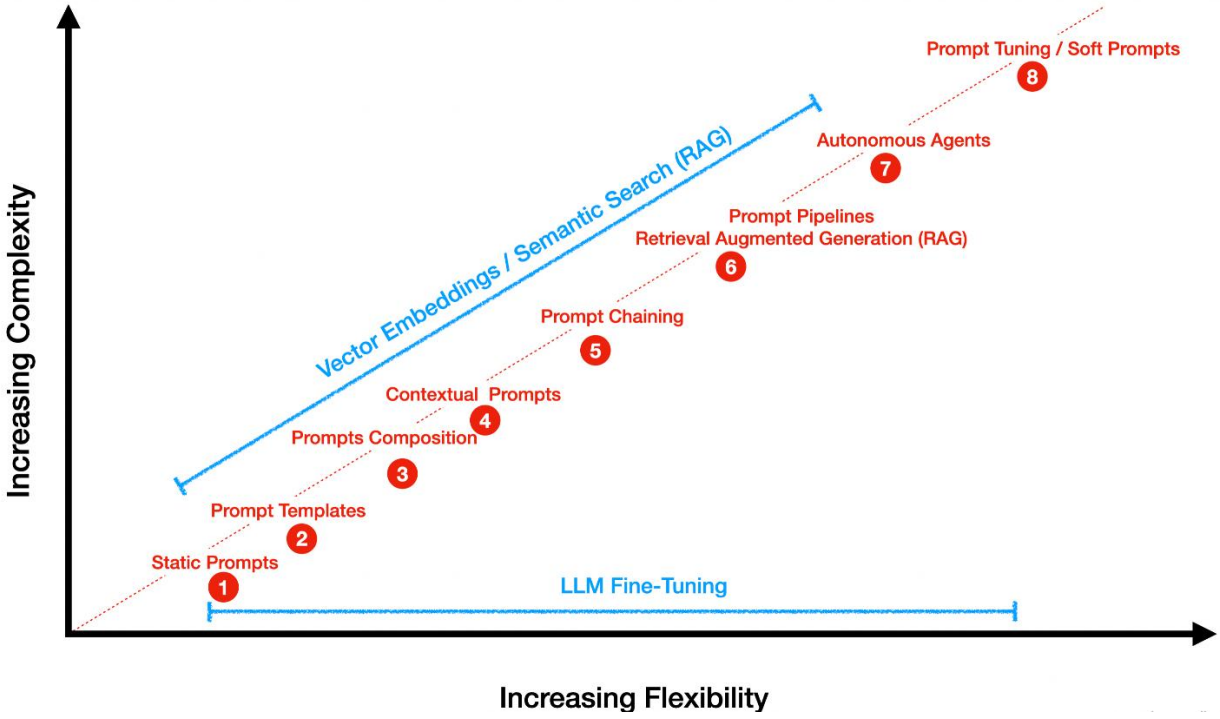
Table of Contents

1. Introduction

Generative AI has shifted the paradigm of computing, instead of coding explicit instructions, we now instruct models through natural language as in the language of our daily use be it Hindi, English, Arabic or any other languages of general use. The quality of these instructions technically we call it prompts, directly governs the reliability, interpretability, and usability of model outputs. Small variations in phrasing can result in vastly different responses, making prompt writing not just a skill but a science of model control.

Moreover, the effectiveness of prompts is highly dependent on the model or platform being used. For example, ChatGPT (OpenAI) is optimized as a conversational model, excelling at structured and human-like responses. Perplexity AI functions more like a research assistant, tightly integrated with retrieval, thus performing best for knowledge-grounded queries. Claude (Anthropic) is widely recognized for strong reasoning, coding explanations, and ethical alignment, Cohere's models are tuned for enterprise-scale language tasks with automated prompt optimization tools. Even open-source models like LLaMA, Mistral, and Falcon exhibit unique prompt-response behaviours. Thus, *prompt writing cannot be separated from model choice*, as the same instruction can lead to different quality and styles of output across platforms.

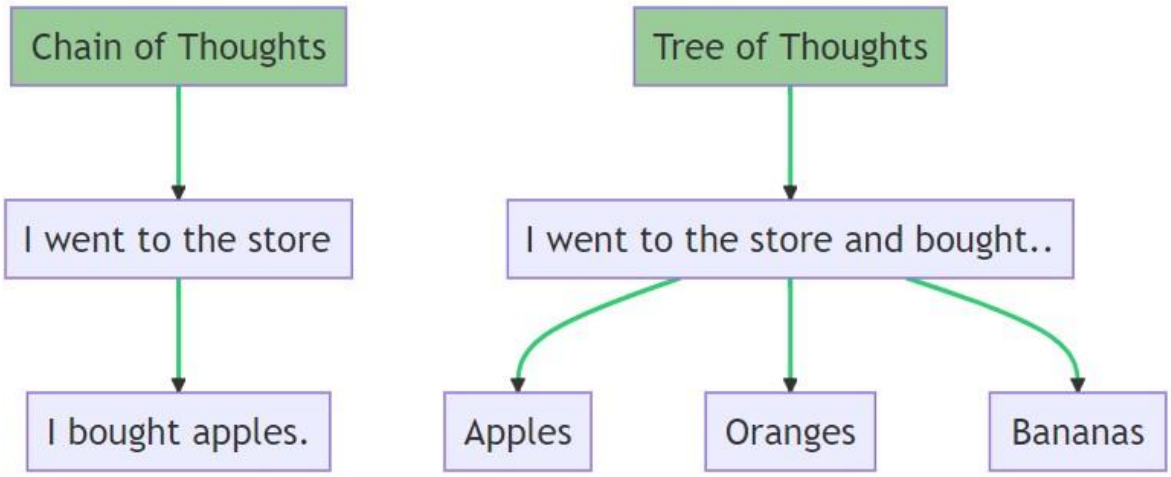
My research surveys the most influential prompt engineering frameworks (2022–2025), compares their strengths, and discusses how they support better model control in real-world applications.



2. Evolution of Prompt Frameworks (2022–2025)

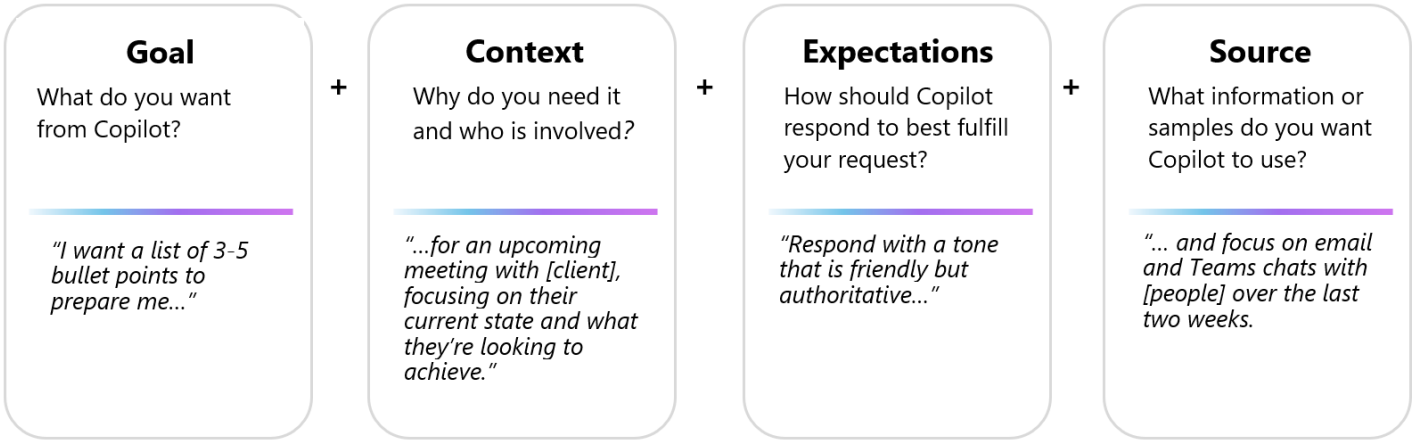
Foundational Research Methods

1. Chain-of-Thought (CoT) Prompting
    - Jason Wei, Google Research - OpenAI, 2022
    - Concept: Instructs the model to reason "step by step" before concluding.
    - Impact: Greatly improves accuracy of reasoning-heavy tasks like logic puzzles, math, and multi-step problem solving.
2. ReAct (Reason + Act) Framework
    - Shunyu Yao, Princeton & Google Brain, 2022
    - Concept: Merges reasoning traces with actions such as tool calls or retrieval.
    - Impact: Allows LLMs to not only think but also interact, making them problem-solving agents.
3. Tree-of-Thought (ToT) Prompting
    - Shunyu Yao, Princeton & Google Brain, 2023
    - Concept: Extends CoT by letting the model explore multiple reasoning branches simultaneously.
    - Impact: Useful for decision-making, optimization, and tasks with multiple possible solutions.



Enterprise & Practitioner Frameworks

4. Anthropic's 10-Step Internal Prompt Template
    - Anthropic (Claude team), 2025
    - Concept: A rigorously structured enterprise framework covering role, context, tone, reasoning, and output formatting.
    - Impact: Ensures reliability and reproducibility in high-stakes deployments.
5. Microsoft Copilot's GCES (Goal Context Expectation Source) Framework
    - Microsoft Copilot team, 2024
    - Concept: Maps user goals to outputs by explicitly stating objectives, context, and references.
    - Impact: Widely adopted in organizational settings for its clarity and reproducibility.
6. The Anatomy of an o1 Prompt
    - Ben Hylak (Co-founder, Raindrop), popularized by Greg Brockman (Co-founder, OpenAI), 2025
    - Concept: Breaks down prompts into simple, modular parts (setup, context, guidance, expected output).
    - Impact: Offers clarity and accessibility for practitioners, particularly in everyday usage of GPT-4o and o1 models.

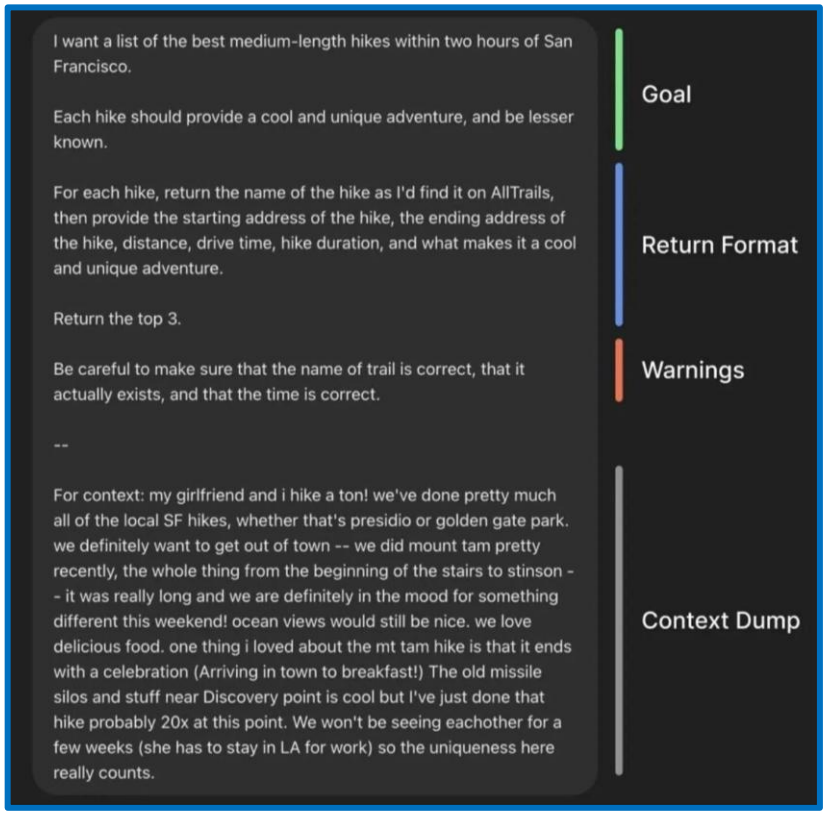


3. Structure of a Good Prompt

A well-designed prompt generally follows six components:

1. Role / Persona → Define the model's perspective.
2. Task / Objective → State clearly what must be done.
3. Context / Input → Provide background, data, or reference.
4. Constraints / Rules → Add limits (word count, tone, scope).
5. Output Format → Specify exact structure (table, list, JSON).
6. Reasoning Instruction (as required) → Ask for step-by-step thinking if needed.

7. Example Prompt:
   You are a policy researcher. Analyse the given household consumption dataset (food, health, education). Identify the top 3 expenditure trends. Provide the answer in ≤150 words, suitable for policymakers, in numbered list format. First, think step by step about the data, then summarize.



The above image refers to The Anatomy of an o1 Prompt, curated by Ben Hylak (Co-founder, Raindrop), later this framework got popularized when Greg Brockman (Co-founder, OpenAI), shared this on his social media accounts & public forums.

4. Putting Constraints on a Model

Constraints ensure model control and reproducibility. They can be embedded directly into prompts:

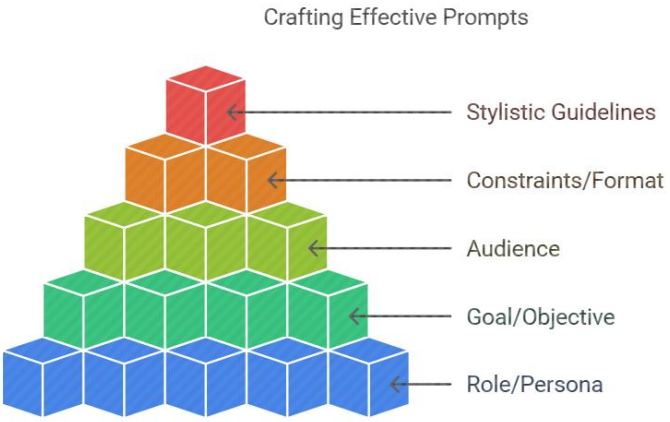
- Output Formatting: Like return in JSON format with keys: Category, Trend, Insight.
- Length & Style: Like explain in ≤200 words for a high-school audience.
- Content Boundaries: As like use only the provided dataset. If unsure, respond with insufficient data.
- Role & Rule Enforcement: Like you are a legal advisor. Do not provide medical or financial advice.

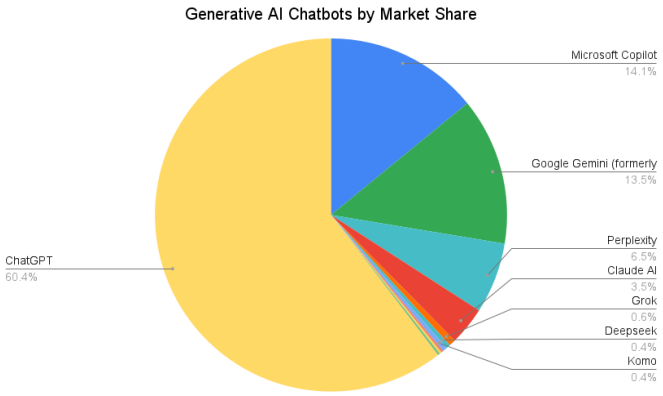
These constraints help prevent irrelevant, hallucinated, or misaligned outputs.

5. Model Dependence in Prompt Effectiveness

Prompt design must account for the strengths of the model:

- ChatGPT (OpenAI) → Best for human like conversational, structured, formatted outputs.
- Claude (Anthropic) → Strong reasoning, long-context analysis, coding tasks.
- Perplexity AI → Retrieval-based research with citations.
- Cohere (Prompt Tuner) → Enterprise automation and empirical optimization.
- Open-source models → Require tighter constraints; high flexibility but less predictable.





6. Comparative Analysis

Framework	Origin	Strength	Weakness	Ideal Use Case
CoT (2022)	Google Research	Strong reasoning accuracy	Verbose outputs	Math, logic, problem solving
ReAct (2022)	Princeton & Google Brain	Combines thought + action	Needs tool integration	Agents, multi-step tasks
ToT (2023)	Princeton & Google Brain	Explores multiple reasoning paths	Computationally expensive	Planning, optimization
Anthropic 10-Step (2025)	Anthropic	Structured, enterprise-grade	Rigid, less creative	Compliance-heavy workflows
Microsoft GCES (2024)	Microsoft	Clear, reproducible outputs	Narrow flexibility	Enterprise copilots, docs
o1 Anatomy (2025)	Raindrop / OpenAI	Simple, modular design	Less research-based	Practical GPT-o1 prompting

7. Discussion: Toward Better Model Control

Prompt engineering has evolved from research prototypes (CoT, ReAct, ToT) to enterprise templates (Anthropic, Microsoft, o1 Anatomy).
For stronger model control, hybrid strategies work best:

- Use CoT/ToT for reasoning scaffolds.
- Combine with ReAct when tool usage is needed.
- Apply Anthropic 10-Step / GCES for reliability and enterprise adoption.
- Simplify with o1 Anatomy for practitioner workflows.

This layered control yields outputs that are accurate, explainable, reproducible, and aligned.

8. Conclusion

Prompt writing has matured into a discipline that blends clarity, structure, and constraints with model-specific awareness. The six frameworks reviewed represent milestones in this evolution.

A good prompt follows a structured format (role → task → context → constraints → output → reasoning). Effective model control comes from embedding constraints and choosing frameworks suited to the model's strengths.

This is the backup research which I have used as a base to write my final prompt for the screening task of FOSSEE Semester Long Internship - Autumn 2025.