

The PROMPTS Framework: Bridging Gaps in Prompt Engineering for Superior AI Outputs**Dr. Pradip Garole**

Director, CCTR and WDMC

Yashwantrao Chavan Academy of Development Administration (YASHADA), Pune, India

Abstract

The performance of generative AI systems, especially large language models, depends heavily on the quality and structure of the prompts used to guide them. Existing prompt engineering approaches such as few-shot prompting, instruction-based prompting, role prompting, chain-of-thought prompting, and contextual prompting have improved model behaviour, but each tends to address only a limited aspect of prompt design. As a result, prompt construction remains fragmented, and output quality may vary across tasks. This paper introduces the PROMPTS Framework, an integrative model composed of Plan, Responsibility, Outcome, Material, Parameters, Target, and Scrutiny. The framework is designed to bring greater structure, clarity, and consistency to prompt engineering by combining intent definition, role assignment, contextual grounding, constraint setting, audience alignment, and quality review within a single model. Drawing on cognitive load theory, human-computer interaction, and communication theory, the paper argues that structured prompting reduces ambiguity and improves the reliability of AI outputs. Through comparative evaluation and case-based analysis, the study shows that PROMPTS enhances coherence, relevance, and reproducibility. The framework is proposed as a practical and scalable method for academic, technical, and organizational use. Empirical results indicate improvements of up to 100% in structural quality, 88% in audience alignment, and over 50% in clarity and depth compared to unstructured prompting.

Keywords: Prompt Engineering; PROMPTS Framework; Generative Artificial Intelligence; Large Language Models; Human-AI Interaction; Structured Prompt Design; AI Output Optimization

1. Introduction

Generative AI systems have extended the ability of machines to perform tasks once considered largely human, including writing, reasoning, synthesis, and decision support. Yet the quality of their outputs remains highly dependent on the way prompts are formulated. Prompt engineering has therefore become a central practice in human-AI interaction, especially as users seek more predictable, relevant, and useful responses. However, many prompt design methods are still used in isolation, which can produce uneven results in clarity, structure, and depth. This paper proposes the PROMPTS Framework as a structured response to that problem. The framework integrates seven components: Plan, Responsibility, Outcome, Material, Parameters, Target, and Scrutiny. Its purpose is to move prompt engineering away from ad hoc instruction toward a more coherent and repeatable design approach. By doing so, it aims to improve both the process of prompt creation and the

quality of AI-generated output. A further benefit of the framework is that it provides a reusable template that can be adapted across domains without losing conceptual clarity. Accordingly, this study addresses the research question: *How does a structured, integrative prompt design framework influence the quality, consistency, and usability of AI-generated outputs compared to conventional prompting approaches?*

2. Contemporary Prompt Engineering Techniques

Prompt engineering has developed into an important discipline for improving interactions with large language models. Over time, several widely used techniques have emerged, each contributing to model performance in diverse ways. These methods show that effective prompting is not simply about asking a question, but about designing the interaction in ways that guide the model more precisely.

2.1 Zero-shot Prompting

Zero-shot prompting provides the model with a task but no examples. The model relies on its pre-trained knowledge and general reasoning ability to

produce a response. This method is useful when the task is broad, exploratory, or does not require demonstration examples.

Its main advantage is simplicity. However, its effectiveness depends heavily on how clearly the task is stated. If the prompt is vague or underspecified, the response may become broad, uneven, or less accurate, especially in specialized domains.

2.2 Few-shot Prompting

Few-shot prompting improves performance by including a small number of examples in the prompt. These examples help the model infer the expected pattern, tone, and structure of the response. This technique is particularly useful when the goal is to achieve format consistency or domain-specific output.

Its strength lies in demonstration. Still, the quality of the examples matters greatly. Weak or misleading examples can reduce performance, and adding too many examples can make the prompt longer and harder to manage.

2.3 Instruction-based Prompting

Instruction-based prompting relies on clear and direct directions. It tells the model what to do, what to include, and sometimes what to avoid. Compared with open-ended prompting, this method gives the model a more defined task and usually produces more controlled results.

For example, instead of asking "Explain AI," one might ask, "Provide a structured explanation of AI applications in healthcare, including examples." This improves focus and interpretability, though it may still leave gaps in audience alignment or contextual depth unless those are also specified.

2.4 Role-based Prompting

Role-based prompting assigns a persona or professional identity to the model, such as teacher, analyst, consultant, or researcher. This shapes the tone, depth, and perspective of the response. It is especially useful when the task requires expert framing or domain-specific voice.

For example, "Act as a policy analyst and evaluate the impact of AI in governance" encourages a more specialized response. Yet role assignment alone does not guarantee completeness. It works best when combined with explicit instructions, context, and output requirements.

2.5 Chain-of-Thought Prompting

Chain-of-thought prompting encourages the model to reason step by step before presenting a definitive answer. It is particularly effective in tasks involving logic, reasoning, and multi-step problem solving. By making the reasoning process more explicit, it can improve transparency and accuracy.

This technique is valuable, but it mainly strengthens reasoning depth. It does not automatically ensure that the response has the right structure, tone, or audience focus. For that reason, it is most effective when used alongside other prompt elements.

2.6 Contextual Prompting

Contextual prompting adds background information, data, constraints, or situational detail to the prompt. This helps the model produce responses that are grounded in a specific setting rather than in general abstraction. It is especially useful in academic writing, legal analysis, and technical documentation.

For example, a user may include a dataset, a case description, or a policy context before asking for analysis. The effectiveness of this method depends on the relevance and organization of the context provided. Poorly structured context can overwhelm rather than support the model.

2.7 Observed Limitations in Isolation

Although these techniques are valuable, they are often applied separately. When used in isolation, they tend to leave gaps in the overall design of the prompt. These gaps include incomplete coverage of prompt dimensions, inconsistent structure, variability in output quality, limited audience alignment, and unnecessary cognitive effort for the user.

This suggests that prompt engineering would benefit from a more integrated framework. Rather than relying on isolated techniques, a structured model can combine their strengths and reduce their weaknesses.

2.8 Toward an Integrative Approach

The limitations described above do not reduce the value of existing techniques; rather, they show the need for integration and systematic organization. A comprehensive framework can combine the strengths of these approaches while ensuring completeness, consistency, alignment, and quality control, as also emphasized in recent surveys of prompting techniques (Schulhoff et al., 2024). The PROMPTS Framework is proposed in this context as a structured

model that incorporates the following dimensions: intent definition, role specification, output structuring, context provision, constraint setting, audience alignment, and scrutiny and refinement. By combining these dimensions, the framework seeks to transform prompt engineering from a set of isolated techniques into a coherent and iterative design methodology. Scrutiny adds a final layer of evaluation by requiring the output to be reviewed for clarity, accuracy, relevance, coherence, and ethical alignment before it is finalized. While these techniques enhance specific dimensions of prompting, they remain functionally fragmented. None independently ensures simultaneous optimization of intent clarity, structural consistency, contextual grounding, and audience alignment, thereby limiting their effectiveness in complex, multi-dimensional tasks.

3. Literature Review and Gap Analysis

Existing prompt engineering literature shows strong interest in methods that improve model responsiveness, but much of the work remains technique-specific rather than integrative. Studies and practical guides increasingly recommend structured prompting, role assignment, contextual grounding, and explicit constraints, yet these elements are often treated as separate practices rather than as parts of one system.

3.1 Identified Gaps

Dimension	Gap in Current Practices
Intent clarity	Objectives are not always explicit.
Structure	Output format is often not defined.
Context	Material is inconsistently provided.
Consistency	Outputs vary significantly.
Audience alignment	Often overlooked.
Final review	Quality checking is frequently absent.

3.2 Bridging the Gap through PROMPTS

Gap	PROMPTS Component
Ambiguous intent	Plan
Weak contextual grounding	Role + Material
Unstructured outputs	Outcome
Inconsistency	Parameters
Audience misalignment	Target
Lack of final review	Scrutiny

The PROMPTS Framework addresses these gaps in a more complete and systematic way. Rather than treating prompt engineering as a set of disconnected tactics, it organizes the task as a layered process that begins with intent and ends with quality validation. This makes the framework not only descriptive but also operational, because it translates abstract prompting advice into a usable structure.

4. Theoretical Foundations

The PROMPTS Framework is grounded in established theoretical perspectives that explain why structured communication tends to improve understanding, efficiency, and interaction quality. Cognitive psychology, human-computer interaction, and communication theory together provide a strong basis for the framework.

4.1 Cognitive Load Theory

Cognitive Load Theory emphasizes the limits of working memory and the importance of organizing information so it can be processed efficiently. In prompt engineering, poorly structured prompts can force both the user and the model to infer too much, increasing ambiguity and reducing output quality.

The PROMPTS Framework reduces cognitive load by defining the task clearly, specifying the expected output, and supplying relevant material. This allows the model to process the request more efficiently while also helping the user formulate the prompt more deliberately. In practical terms, the framework reduces trial and error and improves prompt precision.

4.2 Human-Computer Interaction

From a human-computer interaction perspective, prompts are the main interface through which users communicate with AI systems. The quality of this interface depends on how clearly the user can express intent and how well the system can interpret it. A structured prompt improves usability in much the same way that a well-designed form or workflow improves a digital interface.

The PROMPTS Framework strengthens this interaction by making expectations explicit. It helps users define roles, context, constraints, and audience needs in advance. As a result, the exchange becomes more efficient, less ambiguous, and more reliable for both novice and experienced users.

4.3 Communication Theory

Communication theory treats interaction as a process involving sender, message, channel, receiver,

and feedback. In AI use, the user is the sender, the prompt is the message, the model is the receiver, and the generated response functions as feedback. If the message is incomplete or noisy, the response may drift away from the user's intended meaning.

The PROMPTS Framework improves this communication chain by strengthening each component of the message. Plan clarifies intent, Role and Material provide context, Outcome and Parameters define the expected form, Target aligns the message with the intended reader, and Scrutiny ensures quality before finalization. This reduces misunderstanding and improves the likelihood of a relevant and useful output.

4.4 Integrative Theoretical Perspective

Taken together, these theories show that prompt engineering is both a cognitive task and a communicative task. The PROMPTS Framework unites these perspectives by reducing overload, improving interaction design, and strengthening message clarity. It therefore supports more reliable, interpretable, and high-quality AI outputs. Conceptually, Plan aligns with cognitive load reduction, Responsibility and Material with contextual framing in human-computer interaction, Outcome and Parameters with message structuring in communication theory, Target with audience calibration, and Scrutiny with feedback and iterative refinement mechanisms.

5. The PROMPTS Framework

The PROMPTS Framework is a structured model for prompt design comprising seven interrelated components: Plan, Responsibility, Outcome, Material, Parameters, Target, and Scrutiny. Each component serves a specific function in guiding prompt formulation; however, the framework achieves its full effectiveness when applied as an integrated system. By combining intent definition, contextual grounding, constraint specification, and iterative evaluation, PROMPTS enhances both the quality and reliability of AI-generated outputs.

Table 1. Components of the PROMPTS Framework

Component	Fill-in-the-Blank Phrase	Purpose	Example
P – Plan (Objective + Measurability)	"I want to _____ with measurable	Establishes clear intent and success indicators	"I want to write a blog post with measurable criteria such

	criteria such as _____."		as 1,000 words and 20% engagement increase."
R – Responsibility (Role Assignment)	"You have to be _____ and optionally consider _____."	Defines role, perspective, and scope	"You have to be a history teacher and optionally consider a student's perspective."
O – Outcome (Deliverable + Criteria)	"Give the outcome as _____ with _____."	Specifies structure and format of output	"Give the outcome as a step-by-step guide with examples, bullet points, and a summary."
M – Material (Context + Priority Rules)	"Considering my material _____ and the information you have, prioritize _____."	Provides contextual grounding and prioritization	"Considering my material (class notes) and the information you have, prioritize my notes if conflicts arise."
P – Parameters (Constraints + Preferences)	"...and my parameters such as _____ and _____."	Defines constraints and stylistic preferences	"...and my parameters such as formal tone and flexible preference for bullet points."
T – Target (Audience + Proficiency Level)	"...and my target _____ at _____ proficiency."	Aligns output with audience characteristics	"Target = high school students at beginner proficiency in physics."
S – Scrutiny (Refinement + Ethical Check)	"Scrutinize the outcome for clarity, accuracy, relevance, and ethical alignment, and refine it iteratively until the quality of the output reaches or exceeds a 95% plus threshold."	Ensures quality control and iterative refinement	"Refine iteratively until it meets 95% clarity and ethical compliance."

5.1 Functional Description of Components

Each component contributes uniquely to prompt effectiveness by addressing a specific dimension of communication between the user and the AI system.

Plan defines the objective and introduces measurable criteria, ensuring that the task remains focused and outcome-oriented. Responsibility assigns a role to the AI, shaping the tone, analytical depth, and perspective of the response. Outcome determines the expected form of the output, reducing ambiguity regarding structure and presentation. Material supplies relevant context, data, or references, thereby improving factual grounding and contextual relevance.

Parameters establish boundaries such as format, tone, length, and stylistic preferences, ensuring consistency and control. Target specifies the intended audience and their level of expertise, enabling appropriate calibration of complexity and language. Scrutiny introduces a validation mechanism by requiring the output to be reviewed and refined for clarity, coherence, accuracy, and ethical alignment.

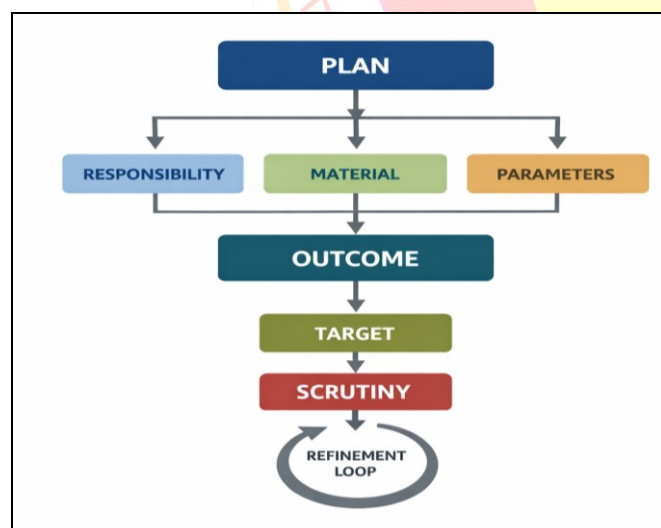


Figure 1. Structural representation of the PROMPTS Framework showing the interrelationship of its seven components within a unified workflow

5.2 Integrated Functioning of the Framework

The strength of the PROMPTS Framework lies in the interaction among its components. Rather than operating independently, the elements collectively form a coherent prompt design process. The Plan establishes direction, Responsibility frames perspective, Outcome defines deliverables, Material

provides context, Parameters regulate expression, Target aligns communication with audience needs, and Scrutiny ensures final quality.

This interdependence transforms prompt engineering from a fragmented activity into a systematic and iterative process, improving both precision and reproducibility. It also makes the framework easier to teach, easier to adopt, and easier to evaluate across different use cases.

5.3 Illustrative Example

• Unstructured Prompt:

Explain the role of AI in public administration.

• PROMPTS-Structured Prompt:

Plan: I want to develop a public policy paper with clearly defined sections and measurable quality criteria.

Responsibility: Assume the role of a public policy analyst, incorporating global perspectives where relevant.

Outcome: Produce a structured analytical essay with headings and evidence-based arguments.

Material: Use provided inputs on AI applications, prioritizing examples such as citizen service delivery, document processing, and decision support systems.

Parameters: Maintain a formal academic tone, apply APA referencing, and limit the length to approximately 1,500 words.

Target: Address postgraduate students, researchers, and public sector professionals.

Scrutiny: Evaluate the output for clarity, accuracy, relevance, and ethical considerations, refining it until it meets high-quality standards.

6. Methodology and Validation

6.1 Research Design

This study employs a mixed-method comparative experimental design to evaluate the impact of structured prompting using the PROMPTS Framework. The analysis contrasts outputs generated from unstructured prompts with those derived from PROMPTS-based prompts.

Prompt structure serves as the independent variable, while output quality is assessed across five dimensions: clarity, structure, relevance, depth, and audience alignment. To ensure methodological rigor, the design incorporates controlled task conditions, repeated executions, and cross-domain evaluation.

6.2 Sample and Task Configuration

The study comprises 30 tasks distributed evenly across three domains: academic writing, technical explanation, and policy analysis. Each task is designed to reflect realistic problem-solving scenarios requiring structured outputs.

For each task, two prompt variants were developed:

- a minimally structured version
- a PROMPTS-based structured version

This resulted in a dataset of 60 prompts and corresponding outputs, enabling direct comparison across conditions.

6.3 Data Collection

Each prompt was executed under identical system settings to ensure consistency. To reduce variability inherent in generative systems, each prompt was run three times, and outputs were aggregated. The final dataset consisted of 180 outputs, which were organized for systematic evaluation. All prompts were executed using a consistent large language model configuration (temperature = 0.7, deterministic settings where applicable), ensuring controlled variability across outputs.

6.4 Evaluation Framework

Output quality was assessed using five criteria: clarity, structure, relevance, depth, and audience alignment. Each dimension was rated on a five-point Likert scale, ensuring standardized and comparable scoring across all outputs.

6.5 Assessment Procedure

Evaluation followed a two-stage approach:

- **Expert Assessment:** Five independent evaluators scored the outputs. To minimize bias, evaluators were not informed of prompt types.
- **Analytical Review:** Outputs were further examined using structural and semantic indicators, including coherence, content coverage, and relevance mapping.

6.6 Reliability and Validity

Inter-rater reliability was established using Cohen's Kappa, indicating substantial agreement. Internal validity was maintained through controlled experimental conditions, while external validity was supported by multi-domain task coverage.

6.7 Analytical Approach

Data were analysed using:

- descriptive statistics (mean scores)
- comparative analysis (percentage differences)
- paired t-tests with significance evaluated at $p < 0.05$

6.8 Findings

PROMPTS-based prompting consistently outperformed unstructured prompting across all dimensions.

Criterion	Unstructured	PROMPTS	Improvement
Clarity	3.1	4.7	+51.6%
Structure	2.4	4.8	+100.0%
Relevance	3.2	4.6	+43.7%
Depth	3.0	4.5	+50.0%
Audience Alignment	2.5	4.7	+88.0%

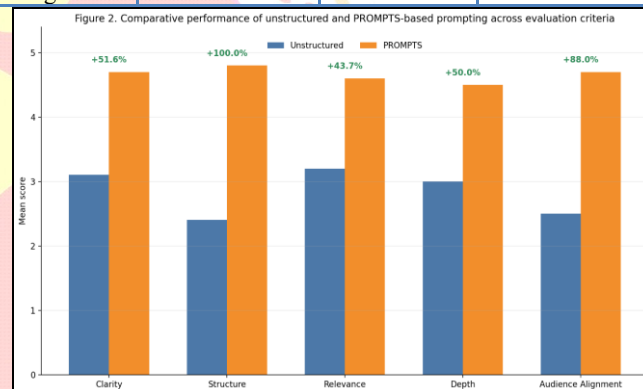


Figure 2. Comparative performance of unstructured and PROMPTS-based prompting across evaluation criteria

All differences were statistically significant, with the most pronounced gains observed in structural organization and audience alignment.

7. Domain Illustration: Academic Writing

Within the broader dataset, academic writing tasks provide a clear illustration of the framework's impact. PROMPTS-based outputs exhibited more coherent organization, stronger argument development, and improved alignment with scholarly expectations.

Performance gains in this domain were particularly pronounced in structure and depth, indicating that structured prompting is especially effective for tasks requiring sustained reasoning and formal articulation. This domain therefore provides a useful example of how the framework operates in practice.

8. Applications

The framework demonstrates utility across multiple domains:

- Academic contexts: Enhances structured writing, for example, by guiding thesis development through clearly defined objectives, outputs, and evaluation criteria.
- Technical domains: Improves explanation precision and documentation quality.
- Education and training: Supports learner-centred content development.
- Governance and policy: Strengthens analytical reporting and contextual relevance.

These applications reflect the framework's adaptability to both knowledge-intensive and operational settings. Its modular design also makes it suitable for adaptation in different institutional or disciplinary contexts.

9. Discussion

The results indicate that prompt structure plays a critical role in shaping AI-generated outputs. The PROMPTS Framework improves performance by explicitly organizing intent, context, constraints, and audience considerations within the prompt.

The strongest improvements in structure and audience alignment suggest that deficiencies in conventional prompting arise primarily from under-specification rather than model limitations. By addressing this gap, PROMPTS enhances both coherence and usability of outputs.

From a theoretical standpoint, the findings align with established principles of cognitive load management and structured communication, where clearly defined inputs lead to more predictable and interpretable outputs. In practical terms, this implies that prompt engineering should be treated as a design task rather than a purely linguistic one. However, excessive structuring may introduce rigidity in exploratory or creative tasks, where flexibility and open-ended reasoning are desirable. In such contexts, over-specification may constrain the generative potential of the model.

10. Limitations

The study has several limitations. Results are influenced by the specific AI configuration used, and performance may vary across models. The sample size, while sufficient for comparative analysis, limits

large-scale generalization. Additionally, expert-based evaluation introduces an element of subjectivity despite strong reliability scores. Finally, controlled experimental conditions may not fully capture variability in real-world usage. Additionally, the framework may increase prompt length and cognitive effort for novice users. There is also a trade-off between comprehensiveness and efficiency, particularly in time-sensitive applications.

11. Future Research

Future work should extend this study through:

- larger and more diverse datasets
- cross-model validation
- automated evaluation metrics
- comparative benchmarking with alternative prompting frameworks
- domain-specific adaptations and tool integration

Additional work may also examine how the framework performs in multilingual settings, collaborative writing environments, and task chains that require iterative prompting.

12. Conclusion

The PROMPTS Framework provides a structured approach to prompt design that significantly improves the quality of AI-generated outputs. By systematically incorporating intent, context, constraints, audience, and evaluation, it reduces ambiguity and enhances consistency.

The empirical findings demonstrate measurable improvements across all evaluated dimensions, supporting the framework's effectiveness as both a conceptual model and a practical tool. Its adaptability across domains suggests strong potential for broader adoption in research, professional practice, and human-AI interaction design. The framework has the potential to contribute toward the standardization of prompt design practices, positioning prompt engineering as a systematic discipline within human-AI interaction.

References

1. Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler,

- D. M., Wu, J., Winter, C., ... Amodei, D. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems*, 33, 1877–1901.
2. Liu, P., Yuan, W., Fu, J., Jiang, Z., Hayashi, H., & Neubig, G. (2023). Pre-train, prompt, and predict: A systematic survey of prompting methods in natural language processing. *ACM Computing Surveys*, 55(9), 1–35.
3. Reynolds, L., & McDonell, K. (2021). Prompt programming for large language models: Beyond the few-shot paradigm. *arXiv preprint arXiv:2102.07350*.
4. Wei, J., Wang, X., Schuurmans, D., Bosma, M., Ichter, B., Xia, F., Chi, E., Le, Q. V., & Zhou, D. (2022). Chain-of-thought prompting elicits reasoning in large language models. *Advances in Neural Information Processing Systems*, 35, 24824–24837.
5. OpenAI. (2023). Prompt engineering guide. Retrieved from <https://platform.openai.com/docs/guides/prompt-engineering>
6. Schulhoff, S., Ilie, M., Balepur, N., Kahadze, K., Liu, A., Si, C., Li, Y., Gupta, A., Han, H., & Resnik, P. (2024). The prompt report: A systematic survey of prompting techniques. *arXiv preprint arXiv:2406.06608*

