Prompting as an emerging skill for Healthcare Professionals

Michele Carmone [1], Leonardo Priore [2] and Antonio Alemanno [1],*

1. *Fisica Sanitaria – AOU Policlinico Foggia;*
2. *Pronto Soccorso – AOU Policlinico Foggia*

Corresponding author: aalemanno@ospedaliriunitifoggia.it

KeyWords:
Prompt engineering; Generative AI, Healthcare professionals

Abstract

Generative Artificial Intelligence (AI) models, such as Large Language Models (LLMs), are opening new avenues in the healthcare sector. These technologies can facilitate clinical documentation, enhance pa-tient education, improve professional training, and streamline the review of scientific literature. However, they may occasionally produce "hallucinations," i.e., responses not grounded in evidence and potentially misleading. In this context, prompt engineering—the careful and stra-tegic design of instructions provided to these models—plays a critical role in guiding outputs, reducing errors, and ensuring greater reliability. This article examines the importance of prompt engineering in healthcare, describing the iterative process, fundamental principles, and various prompting strategies (zero-shot, few-shot, chain-of-thought, au-to-consistency). It further outlines practical applications, addresses ethical challenges, and discusses fu-ture perspectives in integrating this approach into healthcare practice and research.

Background

The generative artificial intelligence (AI) is rapidly transforming the healthcare landscape, offering unprecedented potential to improve patient care, optimize clinical workflows, and en-hance medical research [1-2]. As healthcare institutions and medical companies begin integrating generative AI into their activities, an increasing number of healthcare professionals and re-searchers are expected to work with this technology and fully understand its potential [3]. Therefore, healthcare professionals will need to develop a new essential skill: prompt engineer-ing.

In the context of generative AI, prompt engineering refers to the strategic design and optimiza-tion of inputs, commonly known as "prompts," to guide AI models in generating specific and de-sired outputs. The quality of the prompt plays a critical role, directly influencing the accuracy, relevance, and clinical utility of the AI-generated results [4-5].

Considering the vast amount of training data used by generative artificial intelligence models, identifying and retrieving the most accurate and relevant information poses a significant chal-lenge. Additionally, LLMs are susceptible to a phenomenon known as "hallucination," which oc-curs when the model generates outputs that appear evidence-based but, upon closer examina-tion, often contain substantial errors. Employing advanced prompting techniques is therefore crucial to mitigating such hallucinations. In the healthcare context, hallucinations represent one of the major limitations of LLMs, as they can lead to misdiagnoses and the dissemination of inac-curate information, erroneously presented as factual [6-7].

Prompt engineering process

The process of prompt engineering involves:

- **Clear definition of the objective and desired output:**

identify the purpose of the prompt from the outset, such as synthesizing clinical data, formulating diagnoses, or extracting specific medical information. Clearly specify the type of response required, such as text, tables, or lists, while considering the challenges and nuances associated with the relevant healthcare topic [8].

> For example: Create an Excel table with a weekly schedule of radiological exams, including the day, time slot, type of exam, patient identifier, and clinical priority ('Urgent,' 'Routine,' 'Post-Operative Check-Up'). Adhere to the technical times (30 minutes with contrast, 15 minutes without) and assign urgent priorities to the earliest time slots, evenly distributing other exams throughout the week.

- **Contextualization and domain knowledge:**

base the prompt on solid evidence-based knowledge, aligning it with clinical guide-lines and sector-specific scientific literature. Integrate terminology, protocols, and specialized concepts to guide the model to-wards consistent, accurate, and contextually relevant responses for the healthcare domain [9].

JAHC | Journal of Advanced Health Care
Official Journal of TSRM PSTRP
ESSAY

CITATION:
M. Carmone et al. "Prompting as an emerging skill for healthcare profes-sionals"
JAHC Essay 2025

Received: 2024-12-13
Revised: 2024-12-23
Accepted: 2025-01-20
Published: 2025-01-21

For example: You are a dietitian tasked with creating a weekly meal plan for a 50-year-old man with type 2 diabetes. Consider the following:
Guidelines: Follow the 2024 American Diabetes Association (ADA) nutritional recommendations, which include a balanced diet with 45-60% of daily calories from carbohydrates, healthy fats, and lean proteins.
Nutritional Requirements: Distribute 1,800 kcal per day into three main meals and two snacks. Prioritize low-glycemic index carbohydrates, such as whole grains and legumes, and avoid refined sugars.
Food Preferences: The patient prefers simple, quick-to-prepare meals. He avoids red meat but enjoys fish and tofu.
Goal: Manage blood glucose levels and improve body weight while enhancing insulin sensitivity.
Required Format: Provide the plan in a daily table format, specifying breakfast, lunch, dinner, and snacks. Include precise portions and a weekly shopping list.
Additional Instructions: Add practical recommendations for meal preparation and suggest food substitutions (e.g., oats instead of packaged cereals).

- Incremental approach and choice of prompt type: start with a simple prompt, such as the zero-shot approach, where no exam-ples are provided to the model [10-11]. This method is useful for obtaining quick responses to straightforward questions and is particularly effective in the training of healthcare professionals [12]. Gradually add details, instructions, and constraints. The use of examples, known as few-shot prompting, can further illustrate the intended task and help the model learn more quickly.
Chain-of-thought prompting or self-consistency to progressively increase complexity and ensure greater control over the quality of the output [13-14-15] (Table 1).

Prompt	Examples
Zero-shot The technique involves asking a language model to per-form a specific task without providing any explanatory examples in the prompt, relying solely on the model's general understanding.	A 70-year-old patient with a history of hypertension and type 2 diabetes pre-sents to the emergency department complaining of chest pain and shortness of breath. Describe the priority nursing actions you would take in this situa-tion.
Few-shot In this technique, a few explanatory examples (usually one to five) are provided within the prompt to guide the model in performing a specific task, without requiring additional training.	Analyze the results in the provided file and assign an urgency level (High, Medium, Low) for each blood parameter. **For example:** 1. A patient with hemoglobin at 6 g/dL is classified as High urgency. 2. A patient with hemoglobin at 12 g/dL is classified as Low urgency.
Chain-of-thought It encourages the model to explicitly articulate, step by step, the logical reasoning required to solve a problem or complete a task. This strategy helps improve the con-sistency and accuracy of responses, especially for tasks requiring complex or multi-step decision-making process-es.	A 6-year-old child has difficulty pronouncing sounds like 'r' and 's'. Provide an intervention plan, articulating **step by step** the reasoning behind your rec-ommendations.
Self-consistency This prompt is an advanced prompting technique for lan-guage models that relies on generating multiple respons-es to the same input, followed by identifying the most consistent or frequent response among those generated. This strategy enhances the reliability of the result by se-lecting the most representative or logical option, mitigat-ing uncertainty or inconsistencies in individual outputs.	You are an expert nurse specializing in multiple and independent clinical evaluations. For each clinical scenario I present to you, you will need to gen-erate several reasoned responses and subsequently analyze them to identify the most coherent and reliable solution. Evaluation Procedure: 1. Generate 3-5 independent approaches to the clinical situation. 2. Compare the responses by identifying: - Convergences among the proposed solutions. - Significant discrepancies. - The most solid clinical rationale. Evaluation Criteria: - Patient safety. - Clinical evidence. - Professional guidelines. Scenario Example: An elderly patient with pneumonia, persistent fever, and oxygen saturation fluctuating between 89-92%.

ESSAY
JAHC | JOURNAL OF ADVANCED HEALTH CARE
OFFICIAL JOURNAL OF TSRM PSTRP

- **Iterative refinement, continuous evaluation, and improvement:**

prompting is an iterative process that requires continuous re-finement, systematic evaluation against clinically relevant performance metrics, and constant adaptation based on feedback from healthcare professionals [16]. By testing initial prompts, analyzing outputs, and making iterative revisions, healthcare professionals can progressively improve accuracy, utility, and alignment with the original objectives, ensuring that generative AI models pro-duce reliable, pertinent results that effectively address the needs of patients and other professionals [17-18] (Figure 1).

validity of the generated outputs [22]. Even with a carefully designed prompt, if the AI does not clearly articulate the logic behind its recommendations (for example, the reasoning for excluding certain conditions), it becomes challenging for healthcare professionals to trust the result and validate its accuracy. By fostering transparency—such as providing detailed explanations and the sources of data used—AI can become a more reliable and ethical tool to support clinicians in their decision-making.

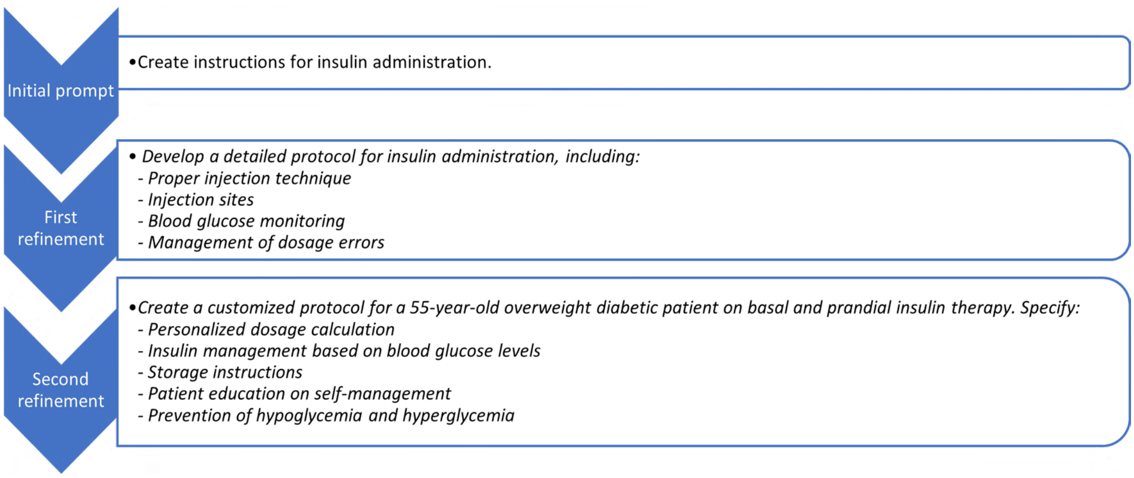

Figure 1: *iterative process: At each step, the prompt becomes more specific, contextualized, and detailed, progressively improving the quality and accuracy of the generated information. The goal is to develop increasingly refined support tailored to the concrete needs of both the patient and the healthcare professional.*

Ethical considerations and inclusivity

In the field of prompt engineering for AI in healthcare, ethical management and bias mitigation are crucial aspects. The use of AI in medical decision-making involves significant risks, including the potential perpetuation or amplification of existing healthcare disparities [19]. To address these risks, it is essential to design prompts that ensure inclusivity, diversity, and neutrality, eliminating any potential distortion related to sensitive demographic factors such as race, ethnicity, gender, or socioeconomic status [20-21].
Moreover, large language models (LLMs) respond to prompts without providing a clear explanation of the logic behind their re-sponses.
This poses a challenge for healthcare professionals, who are responsible for evaluating the accuracy and

Conclusions

In conclusion, the precise definition of prompts, combined with the continuous review and refinement of prompt engineering techniques, represents a new and essential skill for healthcare professionals seeking to fully harness the potential of generative artificial intelligence. The ability to guide models toward relevant, accurate, and contextually consistent responses not only helps mitigate the risk of "hallucinations" and the dissemination of incorrect information but also ensures an ethical, inclusive, and evi-dence-based approach. In this way, the use of generative AI becomes an added value, supporting professionals in making more informed clinical decisions and improving the care experience and health outcomes for all patients.

JAHC | Journal of Advanced Health Care
Official Journal of TSRM PSTRP
ESSAY

Citation:
M. Carmone et al. "Prompting as an emerging skill for healthcare profes-sionals"
JAHC Essay 2025

Received: 2024-12-13
Revised: 2024-12-23
Accepted: 2025-01-20
Published: 2025-01-21

References

1. *Patil, R., Heston, T. F., & Bhuse, V.* (2024). *Prompt engineering in healthcare. Electronics*, 13, 2961. *https://doi.org/10.3390/electronics13152961*
2. *Nazi, Z. A., & Peng, W.* (2024). *Large language models in healthcare and medical domain: A review. Informatics*, 11, 57. *https://doi.org/10.3390/informatics11030057*
3. *Meskó, B.* (2023). *Prompt engineering as an important emerging skill for medical professionals: Tutorial. Journal of Medical Internet Research*, 25, *e50638. https://doi.org/10.2196/50638*
4. *Shah, K., Xu, A. Y., Sharma, Y., Daher, M., McDonald, C., Diebo, B. G., & Daniels, A. H.* (2024). *Large language model prompting tech-niques for advancement in clinical medicine. Journal of Clinical Medicine*, 13(17), 5101. *https://doi.org/10.3390/jcm13175101*
5. *Liévin, V., Hother, C. E., Motzfeldt, A. G., & Winther, O.* (2023). *Can large language models reason about medical questions? ArXiv. https://doi.org/10.48550/arXiv.2207.08143*
6. *Alkaissi, H., & McFarlane, S. I.* (2023). *Artificial hallucinations in ChatGPT: Implications in scientific writing. Cureus*, 15, *e35179. https://doi.org/10.7759/cureus.35179*
7. *Robin, E.* (2023). *ChatGPT: These are not hallucinations; they are inventions and falsifications. Schizophrenia*, 9(1), 52. *https://doi.org/10.1038/s41537-023-00379-4*
8. *Shepherd, J., & Griesheimer, D.* (2024). *FAQs: AI and prompt engineering: The future of nursing education and professional develop-ment. American Nurse Journal*, 19(6). *https://doi.org/10.51256/ANJ062414*
9. *Bonfitto, G. R., Roletto, A., Savardi, M., Fasulo, S. V., Catania, D., & Signoroni, A.* (2024). *Harnessing ChatGPT dialogues to address claustrophobia in MRI: A radiographers' education perspective. Radiography* (*London*), 30(3), 737–744. *https://doi.org/10.1016/j.radi.2024.02.015*
10. *Cai, C. J., Wang, H., Wang, M., Agrawal, R. K., & Bero, N.* (2024). *Prompting is all you need: LLMs for systematic review screening. MedRxiv. https://doi.org/10.1101/2024.06.01.24308323*
11. *Zaghir, J., Naguib, M., Bjelogrlic, M., Névéol, A., Tannier, X., & Lovis, C.* (2024). *Prompt engineering paradigms for medical applications: Scoping review. Journal of Medical Internet Research*, 26, *e60501. https://doi.org/10.2196/60501*
12. *Huang, M., & Sangi-Haghpeykar, H.* (2023). *Prompt engineering in medical education. International Medical Education*, 2(3), 198–205. *https://doi.org/10.3390/ime2030019*
13. *Maharjan, J., Garikipati, A., Singh, N. P., Cyrus, L., Sharma, M., Ciobanu, M., Barnes, G., Thapa, R., Mao, Q., & Das, R.* (2024). *Open-MedLM: Prompt engineering can out-perform fine-tuning in medical question-answering with open-source large language models. Sci-entific Reports*, 14(14156). *https://doi.org/10.1038/s41598-024-64827-6*
14. *Wei, J., Wang, X., Schuurmans, D., Bosma, M., Ichter, B., Xia, F., Chi, E., Le, Q., & Zhou, D.* (2022). *Chain-of-thought prompting elicits reasoning in large language models. ArXiv preprint arXiv*:2201.11903. *https://doi.org/10.48550/arXiv.2201.11903*
15. *Sahoo, P., Singh, A. K., Saha, S., Jain, V., Mondal, S., & Chadha, A.* (2024). *A systematic survey of prompt engineering in large language models: Techniques and applications. ArXiv preprint arXiv*:2402.07927. *https://doi.org/10.48550/arXiv.2402.07927*
16. *Maharjan, J., Garikipati, A., Singh, N. P., Cyrus, L., Sharma, M., Ciobanu, M., Barnes, G., Thapa, R., Mao, Q., & Das, R.* (2024). *Open-MedLM: Prompt engineering can out-perform fine-tuning in medical question-answering with open-source large language models. Sci-entific Reports*, 14(14156). *https://doi.org/10.1038/s41598-024-64827-6*
17. *Wei, J., Wang, X., Schuurmans, D., Bosma, M., Ichter, B., Xia, F., Chi, E., Le, Q., & Zhou, D.* (2022). *Chain-of-thought prompting elicits reasoning in large language models. ArXiv. https://doi.org/10.48550/arXiv.2201.11903*
18. *Wang, L., Chen, X., Deng, X., Wen, H., You, M., Liu, W., Li, Q., & Li, J.* (2024). *Prompt engineering in consistency and reliability with the evidence-based guideline for LLMs. npj Digital Medicine*, 7(1), 41. *https://doi.org/10.1038/s41746-024-01029-4*
19. *Hu, Y., Chen, Q., Du, J., Peng, X., Keloth, V. K., Zuo, X., Zhou, Y., Li, Z., Jiang, X., Lu, Z., Roberts, K., & Xu, H.* (2024). *Improving large lan-guage models for clinical named entity recognition via prompt engineering. Journal of the American Medical Informatics Association*, 31(9), 1812–1820. *https://doi.org/10.1093/jamia/ocad259*

ESSAY

JAHC JOURNAL OF ADVANCED HEALTH CARE OFFICIAL JOURNAL OF TSRM PSTRP

20. *Khan, B., Fatima, H., Qureshi, A., Kumar, S., Hanan, A., Hussain, J., & Abdullah, S.* (2023). *Drawbacks of artificial intelligence and their potential solutions in the healthcare sector. Biomedicine & Materials Devices,* 1–8. *https://doi.org/*10.1007/*s*44174-023-00063-2
21. *Işın, Z. C., Fidan, H., Işın, B. T., Işın, E., & Işın, T.* (2024). *Is Prompt Engineering a Profession? International Journal of Artificial Intelli-gence and Applications* (*IJAIA*), 15(3), 29–38. *https://doi.org/*10.5121/*ijaia*.2024.15303
22. *Harrer, S.* (2023). *Attention is not all you need*: *The complicated case of ethically using large language models in healthcare and medi-cine. eBioMedicine,* 90, 104512. *https://doi.org/*10.1016/*j.ebiom*.2023.104512

Conflicts of Interest: *The authors declare no conflict of interest.*

JAHC | JOURNAL OF ADVANCED HEALTH CARE
OFFICIAL JOURNAL OF TSRM PSTRP

ESSAY

CITATION:

M. CARMONE ET AL. "PROMPTING AS AN EMERGING SKILL FOR HEALTHCARE PROFESSIONALS"

JAHC ESSAY 2025

RECEIVED: 2024-12-13
REVISED: 2024-12-23
ACCEPTED: 2025-01-20
PUBLISHED: 2025-01-21