A Scoping Review on Generative AI prompting to optimize the workflow of healthcare professionals in sub-Saharan Africa

Oluwamayowa Ogundaini[1[0000-0002-3243-6082]] and Lisa-Dionne Morris[2[0000-0002-6348-7675]]

[1] AFDA School of Business Innovation and Technology, Cape Town, South Africa
[2] School of Mechanical Engineering, University of Leeds, Leeds, United Kingdom

Abstract. Generative artificial intelligence (AI) applications have enhanced democratization of information to the extent that industry professionals can automate routine tasks, gain insights from complex data and execute tasks more efficiently through generation of text, image, and audio content. Although these applications augment human capabilities, there are concerns about veracity of AI prompting, which results in hallucinations that could have dire consequences on clinical workflow of healthcare professionals. The impact of prompting patterns on optimization of clinical workflows at points-of-care remains nascent with limited evidence especially in healthcare sectors of sub-Saharan Africa. This review explores existing literature on how generative AI prompt engineering optimizes clinical workflow of healthcare professionals by adopting Arksey and O'Malley five-stage scoping review framework to analyze peer-reviewed publications. A comprehensive search strategy was conducted in scholarly databases, including PubMed, IEEE Xplore, and Google Scholar between 2019 and 2025. The study highlights AI prompt engineering strategies, how prompting affects clinical and administrative activities, and how limitations of generative AI prompting could be addressed. Evidence of generative AI prompt engineering are limited in SSA while the Global North and China are the most dominant regions in the discourse. Consultations, clinical decision support, record summaries and documentation, research and prescription recommendations are leading activities in which AI prompting is perceived as most significant. To conclude, this study provides insights for health managers, healthcare professionals, data scientists, ethicists, health IT experts, human-computer interaction practitioners, and researchers on standardizing integration of generative AI use at points-of-care.

Keywords: Generative AI, Prompt Engineering, Large Language Models, Clinical Workflow, Hallucination, Sub-Saharan Africa

1 Introduction

1.1 Background

There has been an evident increase in the integration of generative Artificial Intelligence (AI) tools (e.g., ChatGPT, Claude and Med-PaLM) in professional domains, including finance, education, manufacturing, automobiles, and the healthcare industry [1]. These tools were developed using large language models (LLMs), computer vision models, and algorithms as conversational agents to generate and personalize text, images, and audio-visual content. Particularly in the health sector, these generative AI tools have been reported or demonstrated the potential to improve patient engagement at points-of-care, timely intervention of diagnoses, enhance medical research during treatment planning, and ultimately optimize clinical workflows [2, 3]. As generative AI rapidly evolves, there is an increasing need for healthcare professionals to adapt and acquire new skill competencies to effectively leverage AI technologies. Thus, the integration of generative AI has created a new critical skill, referred to as AI prompting through an emerging discipline known as prompt engineering [4].

Prompt engineering is a strategic and systematic approach for iteratively formulating inputs or instructions to query pretrained generative AI tools in a manner that produces relevant and desired outputs or perform a natural language processing (NLP) task [5, 2]. These inputs are natural human language instructions commonly referred to as *prompts*. A study by Shah et al [6] argued that AI prompting involves understanding the capabilities and limitations of AI models, as well as the nuances of language, and providing examples that can influence the quality, effectiveness, and relevance of AI-generated yet context-specific content [3]. Therefore, mastering AI prompting and prompt designs require a combination of creativity, critical thinking, and domain-specific knowledge to formulate queries that elicit accurate and useful responses from AI tools, while minimizing hallucinations [7]. AI hallucinations are defined as instances where algorithms generate outputs that are not factual, inaccurate, or not grounded when further examined objectively [6]. These instances may include inconsistencies, false statements, or fabrications that lead to misinformation. Therefore, the role of quality prompting and robust verification mechanisms that involve human oversight is essential for AI development and use especially in instances such as points-of-care.

Prompt engineering requires careful attention to several critical factors: (1) the selection of a language that aligns with the expected outcomes of the generative AI application; (2) provision of precise and unambiguous ed complex instructions; (3) an awareness of inherent biases; and (4) iterative prompting techniques [6]. The first consideration relates to the type of vocabulary or phrasing of statements that might influence the generative AI's output. Second, the arrangement of words in a query or prompt should provide a clear context; otherwise, it will constrain the eventual output or response. Hence, prompts in generative AI should be framed in such a way that the output enhances the creativity of the application to its maximum capability while decreasing its hallucination tendencies. Iterative prompting is a form of refining prompts through trial and error. Through iterative prompting, prompters learn to analyze outputs and identify how to adjust prompts accordingly to ensure contextually relevant results. In other words, generative AI prompt engineering requires a balance between analytical thinking, contextual awareness, and creativity, ensuring that the generated output aligns with task requirements and user expectations – contextual value and relevance.

In the healthcare space, AI prompt engineering requires healthcare professionals to develop the ability to comprehend how these algorithms function, as well as the capability to optimize them for various tasks, such as information synthesis, image recognition, screening, and diagnosis [2]. Data interpretation skills enable professionals to extract meaningful insights from vast amounts of data, which is essential for informed decision making. The quality of the prompting and eventual output plays a critical role in directly influencing the behavior, accuracy, relevance, and utility of the AI-generated results at the point-of-care [8]. For instance, well-designed prompts and refinement offer enhanced clinical decision support by providing accurate diagnostic suggestions or treatment recommendations based on patient historical data; generating health record summaries, which reduces administrative burden [4, 9]. Furthermore, when prompts are well crafted within domain-specific boundaries they can facilitate personalized care plans to create and communicate tailored postoperative care based on a patient's age, comorbidities, and socioeconomic realities [6]. To optimize the workflow of healthcare professionals, context-specific instructions would improve effectiveness of AI-assisted workflows at points-of-care, leading to improved patient-provider experience.

The integration of generative AI into healthcare settings has been accompanied by unprecedented challenges [7, 10]. These challenges include inherent biases in the training data, constraints on quality datasets, transparency and explainability issues, contextual relevance, security and privacy vulnerabilities, and disparities in digital literacy [2, 11]. Consequently, the quality of output, irrespective of prompt engineering and fine-tuning, is less accurate, which is a major limitation of generative AI technologies. Crafting effective prompts to communicate with generative AI often requires iterative tuning, domain expertise, creativity and manual curation of examples which can be time-consuming and inefficient for complex tasks.

Despite well-formulated prompts, internal reasoning mechanisms of generative AI remain largely inaccessible. Consequently, it is difficult to diagnose misinformation, errors, or ensure consistency. In terms of contextual relevance, generative AI prompts may produce a desired outcome in a certain context and fall short in another, particularly when applied across different domains [3]. In other words, generative AI are sensitive to changes in phrasing, formatting, and context. When prompts are based on trial-and-error and not phrased within specific contexts, generative AI tends to generate outputs that are hallucinations. In cases where hybrid prompt engineering strategies are employed while they may improve performance, there are concerns that the prompt structure is complex task for non-domain or non-expert users.

Beyond the challenges associated with information accuracy and explainability of outputs produced by generative AI, there remains a need to address the issues of data quality and availability, which are essential for ensuring the contextual relevance of the generated content peculiar to the health and well-being of the SSA population. Other pertinent challenges relate to equitable access, security, and ethical use of sensitive patient information [2]. These challenges impact healthcare professionals' prompt designs, generated content, task performance, service delivery, and patient outcomes.

1.2 Problem Statement

Hallucinations are a major limitation of generative AI tools powered by large language models (LLMs) and pose a significant risk as they can lead to misdiagnoses, disinformation when used to develop treatment plans, dissemination of false information to patients, and cause significant harm to human health and well-being [5, 12 - 13]. Hence, prompt engineering skills are critical aspects as a fundamental requirement of using AI to support healthcare activities [3, 9]. However, the reported impact of healthcare professionals' generative AI prompting patterns on the optimization of clinical workflows remains limited in sub-Saharan Africa context [7]. Many of the existing scholarly work focuses on the broad adoption of AI.

1.3 Objective of the Scoping Review

The objective of this study was to map the existing literature on how generative AI prompting strategies influence the workflow optimization of healthcare professionals. The goal of this paper is to create awareness of the need for the development of AI prompting skills as an emerging digital competence for healthcare professionals who have (or intends to) adopted AI tools for their work activities. The guiding research question is as follows: *What is the impact of AI prompting on the optimization and outcomes of AI-enabled work activities carried out by healthcare professionals*

1.4 Research Questions

1. What prompting patterns are used in healthcare AI tools?
2. What are the impacts of generative AI prompt engineering on the clinical workflow of healthcare professionals?
3. What are the reported limitations of AI in prompting clinical workflow?
4. How can hallucinations from generative AI prompting and its impact in healthcare settings be minimized to ensure patient safety?

2 Methodology

The study follows Arksey and O'Malley's (2005) five-stage review framework for scoping reviews. This scoping review methodology was employed because it facilitates the exploration and identification of concepts and prevailing topics within a particular research domain or discipline [14, 15]. Scoping reviews are used to unearth and synthesize emerging concepts that may require further investigation for knowledge creation, acquisition, enlightenment and application. For instance, this study sought to establish existing knowledge on prompt engineering of generative AI, how AI-generated content enables task performance, impact of prompt design strategies on optimization of work activities executed by healthcare professionals and how to mitigate hallucinations, especially at points-of-care in SSA. Arksey and O'Malley's five-stage framework for this study is outlined as (1) identification of research questions and objectives, (2) search strategy and eligibility criteria, (3) selection of relevant publications, (4) charting retrieved data to highlight findings, and (5) reporting.

2.1 Identification of Research Questions and Objective(s)

The primary motivation for this study was based on the rapid rate at which generative AI is integrated into daily human activities without a holistic comprehension of managing its impact, which has raised concerns about both the output and risks of disinformation. We argue that there is a need to pay attention to how prompt engineering impacts the crafting of prompts, fine-tuning, and validity of ensuing outputs. In the healthcare industry, the handling of sensitive data and safeguarding human well-being are paramount, necessitating an exceptionally low margin for error, even with the integration of generative AI technologies. Therefore, the objective of this study is to map the existing literature on how generative AI-prompting engineering strategies influence the workflow optimization of healthcare professionals.

2.2 Data Sources and Search Strategy

The search strategy was applied to query peer-reviewed databases, including PubMed, IEEE XPlore, and Google Scholar, to identify existing literature on the intersection between prompt engineering in generative AI and optimization of work activities executed by healthcare professionals. These databases were selected because of their extensive repository of domain-specific content pertaining to prompt engineering in generative AI and health-related research. The keywords used to search the databases were generative AI, LLM, prompt engineering, and healthcare professional. The keywords were combined using the logical operators as follows: (“Generative AI” OR “LLMs”) AND (“prompting” OR “prompt engineering”) AND (“healthcare professionals” OR “clinical workflow”) AND (“sub-Saharan Africa” OR “Africa” OR “Low-and-middle income”). Subsequently, the results were limited to peer-reviewed studies published between 2019 and 2025. The eligibility criteria are described in the next section.

2.3 Inclusion and Exclusion Eligibility Criteria

The inclusion criteria considered only peer-reviewed publications, including journal articles and conference proceedings, within the scope of the study objective. The selected articles must have been written in English because it is the first language and the preferred mode of communication for both authors. The timeline considered was pre-COVID19 till present date due to rapid development and release of generative AI. The exclusion criteria included exemptions from publications written in other languages, closed-access articles, dissertations, theses, and book chapters. The initial filtering process was performed by one author, while the second author verified the extracted data satisfied the objectives of the paper and ensured a consensus of eligible publications. Publications were excluded if they were written in other languages, non-open access, dissertations, theses, or studies conducted outside the specified period, in addition to the SSA context. A summary of the eligibility criteria is provided in **Table 1**.

Table 1. Inclusion and Exclusion criteria for eligible studies selection.

Criterion	Inclusion	Exclusion
Language	English	Non-English
Timeline	2019 – 2025	Before 2019

Topics of interest	Prompt engineering strategies for generative AI use in healthcare settings, impact of AI prompting on healthcare work activities	Use of non-generative AI tools in healthcare settings
Publication types	Journal articles, conference proceedings, book chapters	Dissertations, theses, books and reports
Access	Open access	Closed access
Context	Africa (Especially Sub-Sahara)	Global North

2.4 Study Selection Process

All abstracts were screened based on the eligibility criteria prior to full-text reading and extraction of relevant findings to address the research questions. Subsequently, the selection of relevant publications was based on correlation with the study objective and a thematic analysis as presented in section 3.

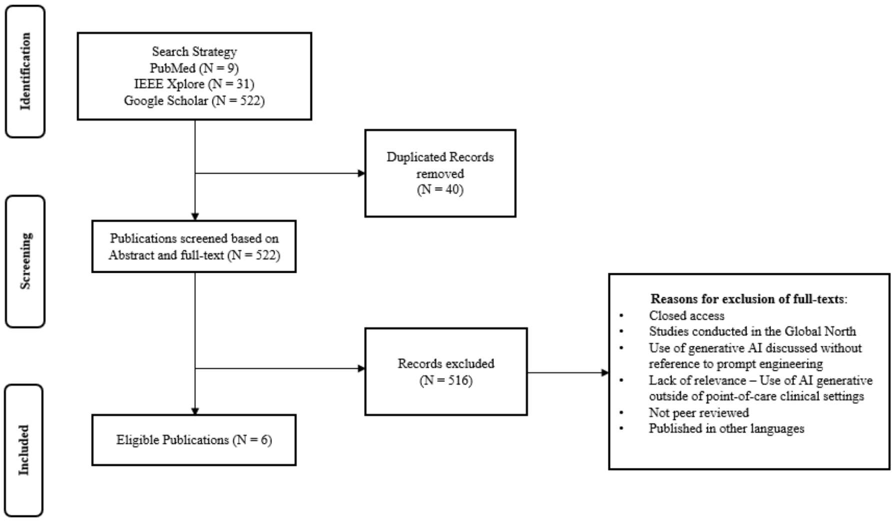

Fig 1. Eligibility criteria for screening publications

2.5 Data Charting Process

Full-text publications that satisfied the inclusion criteria were reviewed in greater detail as illustrated in **Fig *1*** . The authors sorted publications based on title, abstract, and content with contextual relevance to how prompt engineering enabled healthcare professionals to optimize their work activities while utilizing generative AI. Examples of filtered literature include studies on the domain-specific use of generative AI, use of generative AI to accomplish specific health-related tasks, and the evaluation of LLMs based on types of prompt engineering strategies. These studies helped authors identify prompt strategies, impact of generative AI on workflow, and limitations of generative AI, and suggested measures to mitigate concerns such as generative AI hallucinations that may lead to disinformation.

3 Results

The results suggest that scholarly work on AI prompting is emerging and is limited in the Digital Health space of sub-Saharan Africa. Most studies are literature reviews and conceptual papers with a few relating to prompt engineering and its optimization of clinical workflow. The search query on PubMed yielded the least results compared to IEEE Xplore and Google Scholar. Arguable reasons are that IEEE explore offers more technical insights and limited region-specific studies while PubMed addresses clinical applications of generative AI. Google scholar offers a wide range of interdisciplinary journal and conference publications that align to the objective of the study.

With regards to geographical relevance, the results of the query yielded publications that showed that North America, Europe, and China currently dominate the discourse on AI-prompt engineering [7, 11]. The primary areas where prompt engineering is most applicable to optimize the workflow of healthcare professionals can be categorized as diagnosis and treatment, disease surveillance and response, medical imaging and screening, health information documentation and summary, medical research, clinical decision support, and prescription recommendations [7]. These represent the predominant activities for which AI prompt engineering is most pertinent even though evidence indicates that LLMs have the potential to optimize clinical workflows is hindered by systemic, technical, and regulatory challenges.

A descriptive summary is presented in the subsequent section to summarize findings to the research questions. The summary is based on extracting concepts on generative AI applications, healthcare domain or discipline, work activity or task performed (Clinical or Administrative), generative AI tools used, prompt type or strategy, and the reported workflow impact.

3.1 Thematic Categories Identified

Prompting Strategies

The predominant generative AI prompt engineering strategies are chain-of-thought reasoning, zero-shot and few-shot prompting, explicit instruction, ensemble prompting, heuristic methods, self-consistency, and hybrid. Vilakati [16] evaluated how prompt engineering and prompt formulation could be designed to minimize errors and enhance the trustworthiness of responses. The metrics used to assess these prompting strategies were assumption checking, test selection, output completeness, and interpretive quality. According to Venerito et al [5] it is recommended that a combination of few-shot and chain-of-thought prompting could be combined depending on setting and task purpose.

M’gadzah and O’Malley [8] quantified the effect of a complex prompting strategy to assess the diagnostic accuracy of GPT-4 for prevalent ophthalmological conditions by comparing global and low- and middle-income countries (LMIC). Complex prompts yielded a higher accuracy of 90.1% compared to 60.4% for simple prompts. The authors found that complex prompting significantly enhances GPT-4’s diagnostic performance; however, as a diagnostic tool, it requires contextual tailoring to be effective in identifying diagnostic blind spots when dealing with diverse patient populations and specific diseases, especially in the context of LMIC. It is recommended that datasets be expanded to cater for underrepresented conditions to be culturally and clinically adaptive, as well as the need for clinical domain-specific prompt engineering frameworks. Thus, healthcare professionals can contribute to the design of robust prompts with tight constraints and verifiable phrases such that generative AI health tools might hallucinate less and reduce generic diagnoses.

Impacts of generative AI prompt engineering on clinical workflow of healthcare professionals

The findings presented in this section are based on studies in which generative AI was adopted by healthcare professionals to support one or more aspects of their clinical and administrative work activities.

The integration of generative AI into healthcare settings significantly enhances time management where prompting helps shortens patient consultation periods and mitigate against professional burnout or fatigue [17]. For example, through specific prompt designs, generative AI can produce concise summaries of medical reports, which demystifies complex medical language and improves patient comprehension of their health records [7, 9]. Additionally, prompting generative AI aids healthcare professionals to automate the creation of clinical documentation, condensing extensive medical literature, and supporting tasks related to diagnosis, triage, and billing. Traditionally, these tasks have relied on manual methods such as using paper and pen to write or search for health information, which are more time-consuming and labor-intensive than voice-activated generative AI prompt engineering [18].

Reported limitations of generative AI in prompt engineering in clinical workflow

Hallucinations pose significant risks when using generative AI, particularly in the dissemination of misinformation [12]. This issue is especially critical in healthcare settings where the accuracy, integrity, and dependability of data are crucial for making informed decisions related to diagnosis, screening, and treatment planning. Hallucinations stem from the unreliability of algorithms, which may yield inaccurate or non-factual information-based contextual irregularities [3, 8]. Moreover, inconsistencies arise when LLMs produce different responses to the same prompts or variations in the response to changes in the prompt structure. In addition, LLMs may generate biased outputs, reflecting the inherent skewness of their pre-training data, and there are issues concerning the transparency or opacity of the model's decision-making processes.

The existing data governance policies do not consider AI-generated content; hence, there is a likelihood of unfair practices when using an individual's health data to generate AI models and tools. The policies were primarily designed to protect individual privacy and regulate harmful data usage but were obsolete because of rapidly evolving AI technologies [19]. The absence of specific guidelines for AI-generated content creates a regulatory gap that could lead to the exploitation of personal health data without proper oversight measures [5]. This situation could result in a range of unfair practices, including unauthorized data mining and biased algorithms, thereby widening the mistrust of AI products.

Although generative AI tools have the capacity to improve interactions between patients and healthcare professionals, they cannot entirely replace practitioners in the crucial role of fostering trust with patients [7, 20]. Human interaction is indispensable in healthcare to achieve a comprehensive understanding of potential patient issues, necessitating the involvement of healthcare professionals. Consequently, the application of generative AI prompts in healthcare is confined to specific functions that do not eliminate essential human interaction and personal engagement.

Minimization of hallucinations for patient safety

Existing studies recommend measures such as retrieval augmented generation (RAG), prompt engineering strategies, fine-tuning, continuous feedback loops involving back-end adjustments, expert reviews, and user-led testing as some of the current strategies that could be used to evaluate, re-train, validate, and verify AI-generated contents [21]. The scarcity of representative and high-quality datasets from the Global South is recognized as a significant contributor to unintended bias, discrimination, exclusion, and contextually relevant AI-generated content. This challenge is primarily due to most LLMs being developed and trained using datasets predominantly sourced from Global North [8, 11]. This bias in the data sources can lead to hallucinations and out-of-context outputs, particularly when applied to diverse global contexts. Diversifying datasets to train LLMs models can enhance the reliability and cultural sensitivity of generative AI, ensuring that prompts produce outputs that are representative and applicable to different contexts.

4 Discussion of Findings

The findings show that generative AI prompt engineering has a significant impact on the perceived usefulness of technologies to support the execution of point-of-care activities, performance of healthcare professionals, and the quality of informed decision making [4]. In addition, developing an effective AI prompting strategy tends to depend largely on an individual's competency in their field of practice. As generative AI technologies advance in learning, become updated and more sophisticated, the ability to effectively communicate with and leverage these tools inherently becomes an asset in the workplace and beyond.

Prompt engineering influences workflow efficiency and decision support. Owing to the nature of work activities being executed, the purpose of task being performed, and the difference in digital literacy of healthcare professionals, there is a need to standardize prompt templates and evaluation rubrics [6]. In this way, hallucinations are checked, and healthcare professionals can be assured that the feedback from generative AI is near-accurate and dependable for informed decision-making.

The practical implications of issues described in the results section highlights a need for context- and domain-specific prompt engineering guidelines to check interactions between healthcare professionals and generative AI-Digital Health, including summarizing digital health records, diagnoses, screening, and developing treatment plans for patients [2]. If healthcare institutions are to benefit from generative AI and the outputs from prompts especially mitigation of adverse medication risks and avoiding reputation-damaging medicolegal claims from patients, they need to invest in prompt engineering as a strategy to equip their staff on responsible engagement with AI [1]. While rollout of generative AI tools is moving at a rapid pace, sectors such as healthcare need policies that will ensure safe and productive use among professionals rather than them being demotivated and resisting the tools. These policies should be centered on the competence levels of prompt engineering, prompt strategies for generative AI based on domain-specific usage and ensuring that AI developers include both validation and verification measures for AI-generated content.

Some notable limitations of this scoping review include limited studies on how to engineer prompt designs for SSA context, language preference when selecting eligible studies, risk of publication bias, unintended omission of relevant studies, and a lack of expert insight into lived experiences of healthcare professionals who have integrated generative AI into their work processes in SSA healthcare systems.

5 Conclusion

This scoping review highlights how responsible engagement with generative AI could be realized though prompt engineering to ensure near-accurate and reliable outputs. The study provides valuable insights for health managers, prompt engineers, human-computer (HCI) interaction domain practitioners, researchers, and organizations integrating AI into their operations by emphasizing the opportunities in prompt engineering as a digital competence and the need to invest resources in training and feedback. This competence is becoming increasingly important as AI technologies are being integrated into various sectors, including healthcare operations. Moreover, when developing frameworks for the implementation of generative AI, it is crucial to evaluate whether LLM is specifically designed for a particular domain based on pre-trained data. Consideration should be given to natural language processing tasks that can be enhanced, common prompt structures, and implications for subsequent decision-making based on generated content. This strategic approach allows healthcare professionals to consistently develop proficiency in prompt engineering.

Developing proficiency in AI prompting necessitates dedicated training and continuous feedback to refine techniques and adapt to evolving AI capabilities as an essential aspect of personal and professional growth. For example, learning initiatives such as short online professional courses, workshops, and certification programs would play a pivotal role in equipping healthcare professionals with prompt engineering skills. By embracing continuous learning and upskilling, healthcare professionals can ensure that they are well-prepared to navigate the complexities of an AI-enhanced landscape and contribute to the advancement of their respective fields. Generative AI prompting has become a sought-after digital competence, and health stakeholders must invest resources in training professionals to craft effective prompts to generate contextually relevant content that minimizes harmful hallucinations. Also, domain-specific prompt engineers are required to realize the true potential of generative AI.

There is an opportunity for healthcare professionals in the health systems of SSA to develop interdisciplinary skills that match the use of generative AI to support their activities and improve the efficiency of service delivery. Future research is needed to explore novel participatory design approaches on how open-source generative AI can be trained using quality datasets that account for culture-sensitive data and well-crafted domain-specific prompts, to assist community health workers and primary healthcare professionals in optimizing their tasks, when addressing challenges specific to SSA. In essence, collaboration and partnership (Sustainable Development Goals 3, 9, and 17) are needed for benefits of digital transformation to be realized in SSA's health sector.

Acknowledgments.

Not Applicable.

Disclosure of Interests.

The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Ong, J.C.L., Chen, M.H., Ng, N., Elangovan, K., Tan, N.Y.T., Jin, L., Xie, Q., Ting, D.S.W., Rodriguez-Monguio, R., Bates, D.W. and Liu, N.: A scoping review on generative AI and large language models in mitigating medication related harm. *npj Digital Medicine*, *8*(1), p.182. https://doi.org/10.1038/s41746-025-01565-7. (2025)
2. Sai, S., Gaur, A., Sai, R., Chamola, V., Guizani, M., & Rodrigues, J. J.: Generative AI for transformative healthcare: a comprehensive study of emerging models, applications, case studies, and limitations. *IEEE Access*, *12*, 31078-31106. https://doi.org/10.1109/ACCESS.2024.3367715. (2024)
3. Wang, D., & Zhang, S.: Large language models in medical and healthcare fields: applications, advances, and challenges. *Artificial intelligence review*, *57*(11), 299. https://doi.org/10.1007/s10462-024-10921-0. (2024)
4. Meskó, B.: Prompt engineering as an important emerging skill for medical professionals: tutorial. *Journal of medical Internet research*, *25*, p.e50638. https://doi.org/10.2196/50638. (2023)
5. Venerito, V., Lalwani, D., Del Vescovo, S., Iannone, F., & Gupta, L.: Prompt engineering: the next big skill in rheumatology research. *International Journal of Rheumatic Diseases*, *27*(5), e15157. https://doi.org/10.1111/1756-185X.15157. (2024)
6. Shah, K., Xu, A. Y., Sharma, Y., Daher, M., McDonald, C., Diebo, B. G., & Daniels, A. H.: Large language model prompting techniques for advancement in clinical medicine. *Journal of Clinical Medicine*, *13*(17), 5101. https://doi.org/10.3390/jcm13175101. (2024)
7. Moodley, K.: Artificial intelligence (AI) or augmented intelligence? How big data and AI are transforming healthcare: Challenges and opportunities. *South African General Practitioner*, *5*(4), 170-175. https://hdl.handle.net/10520/ejc-mp_sagp_v5_n4_a6. (2024)
8. M'gadzah, S. A. T., & O'Malley, A.: Enhancing Diagnostic Accuracy of Ophthalmological Conditions With Complex Prompts in GPT-4: Comparative Analysis of Global and Low- and Middle-Income Country (LMIC)–Specific Pathologies. *JMIR Formative Research*, *9*(1), e64986. https://doi.org/10.2196/64986. (2025)
9. Yao, Z., Jaafar, A., Wang, B., Yang, Z. and Yu, H.: Do Clinicians Know How to Prompt? The Need for Automatic Prompt Optimization Help in Clinical Note Generation. In *Proceedings of the 23rd Workshop on Biomedical Natural Language Processing* (pp. 182-201). https://doi.org/10.18653/v1/2024.bionlp-1.15. (2024)
10. Attai, K., Ekpenyong, M., Amannah, C., Asuquo, D., Ajuga, P., Obot, O., Johnson, E., John, A., Maduka, O., Akwaowo, C. & Uzoka, F. M.: Enhancing the interpretability of malaria and typhoid diagnosis with explainable AI and large language models. *Tropical Medicine and Infectious Disease*, *9*(9), 216. https://doi.org/10.3390/tropicalmed9090216. (2024)
11. Hussain, S. A., Bresnahan, M., & Zhuang, J.: Can artificial intelligence revolutionize healthcare in the Global South? A scoping review of opportunities and challenges. *Digital health*, *11*, 20552076251348024. https://doi.org/10.1177/20552076251348024. (2025)
12. Twinomurinzi, H., & Gumbo, S.: Generative AI: Concerns, usage, challenges, opportunities and sentiments. *South African Computer Journal*, *37*(1), 120-145. https://hdl.handle.net/10520/ejc-comp_v37_n1_a7. (2025)
13. El Mir, A., Luoga, L.T., Chen, B., Hanif, M.A.& Shafique, M.: Advancing Healthcare in Low-Resource Environments Through an Optimization and Deployment Framework for Medical Multimodal Large Language Models. In *2024 IEEE EMBS International Conference on Biomedical and Health Informatics (BHI)* (pp. 1-8). IEEE. https://doi.org/10.1109/BHI62660.2024.10913565. (2024)

14. Colquhoun, H.L., Levac, D., O'Brien, K.K., Straus, S., Tricco, A.C., Perrier, L., Kastner, M. and Moher, D.: Scoping reviews: time for clarity in definition, methods, and reporting. *Journal of clinical epidemiology*, *67*(12), pp.1291-1294. https://doi.org/10.1016/j.jclinepi.2014.03.013. (2014)
15. Tricco, A.C., Lillie, E., Zarin, W., O'brien, K., Colquhoun, H., Kastner, M., Levac, D., Ng, C., Sharpe, J.P., Wilson, K. and Kenny, M.: A scoping review on the conduct and reporting of scoping reviews. *BMC medical research methodology*, *16*(1), p.15. https://doi.org/10.1186/s12874-016-0116-4. (2016)
16. Vilakati, S.: Prompt Enginneering for Accurate Statistical Reasoning with Large Language Models in Medical Research. (2025)
17. Korom, R., Kiptinness, S., Adan, N., Said, K., Ithuli, C., Rotich, O., Kimani, B., King'ori, I., Kamau, S., Atemba, E. and Aden, M.: AI-based Clinical Decision Support for Primary Care: A Real-World Study. *arXiv preprint arXiv:2507.16947*. (2025)
18. Ogundaini, O. O., de la Harpe, R., & McLean, N.: Integration of M-Health information and communication technologies into the clinical settings of hospitals in sub-Saharan Africa: qualitative study. *JMIR mHealth and uHealth*, *9*(10), e26358. https://doi.org/10.2196/26358 (2021)
19. Paul, L., & Rena, R.: Generative AI in South African Healthcare: Navigating Challenges and Harnessing Opportunities for Industry and Research. *Management Dynamics*, *25*(1), 6. https://doi.org/10.57198/2583-4932.1353. (2025)
20. Er-Rays, Y., & M'dioud, M.: ChatGPT in healthcare facilities: an overview and innovations in technical efficiency analysis. *Available at SSRN 4771070*. https://dx.doi.org/10.2139/ssrn.4771070. (2024)
21. Igwe, K., & Durrhiem, K.: A scoping review of culturally sensitive large language models-based cognitive behavioural therapy for anxiety and depression: Global lessons for African implementation. *Interdisciplinary Journal of Sociality Studies*, *5*(1), a06-a06. (2025)