

Unlocking the potential of AI chatbots

BY IVO DUKIC

The landscape of artificial intelligence (AI) is evolving rapidly, and one of its most exciting domains is large language models (LLMs) powering AI chatbots. These models hold immense potential to transform how we interact with technology, offering personalised assistance in diverse settings, including healthcare. For surgeons and allied health professionals, LLMs present a new frontier, brimming with possibilities to enhance clinical practice, research, and communication.

Feature	Gemini	Copilot	ChatGPT
Developer	Google AI	Microsoft / OpenAI	OpenAI
Underlying Model	Gemini	ChatGPT4	ChatGPT3.5/4
Strengths	Superior reasoning and multi-modal (text, audio, image and video) inputs	Real-world integration through Bing search, image generation, personalised learning potential	Creative writing, engaging storytelling, humour, popular culture references
Weaknesses	May generate less hallucinations, and more accurate answers, less integration into plugins and platforms (released Dec 2023)	Relies on search and may be prone to hallucinations and prone to bias, factual inaccuracies, generates sensational content, less integration into platforms	Hallucinations and prone to bias, factual inaccuracies, generates sensational content
Accessibility	Access through google accounts	Integrated into Bing search and Microsoft Edge browser, limited free access	Free access ChatGPT3.5, \$20 per month for ChatGPT4
Cost	Free for research purposes, paid access for commercial use	Free for basic usage, paid tiers for advanced features	Free access to ChatGPT3.5, \$20 per month for ChatGPT4
Bias	Focus on factual accuracy and neutrality can lead to lack of diverse perspectives	Reliance on Bing search results can reflect biases present in the web	Prone to biases, and potentially offensive outputs
Common Sense	Limited ability to reason and understand real-world context (slightly better than GPT4)	Access to Bing search can provide some real-world context, but overall understanding limited	Lacks common sense reasoning, can misinterpret or make illogical statements

What are LLMs?

Think of LLMs as sophisticated algorithms trained on massive datasets of text and code. This empowers them to understand, summarise, and even generate language, mimicking natural language processing (NLP) used in applications like internet search, email filters, and smart assistants. LLMs, however, take NLP to a whole new level, boasting human-like responses to a wide range of inputs. LLMs can also respond in up to 95 human languages and respond with computer code in a number of languages.

Comparing AI Chatbots: Google Gemini, Copilot and ChatGPT

Three of the leading AI chatbots in the market are Gemini, Copilot and ChatGPT. Each model possesses unique strengths and limitations, often sharing certain vulnerabilities which are summarised in the table above.

The LLM Achilles' Heel: hallucinations and bias

Despite their remarkable capabilities, LLMs have an Achilles' heel: hallucinations. Like an overzealous assistant, they may fabricate information not present in their training data. Additionally, their responses can be influenced by inherent biases within the data they are trained on. Awareness of these limitations is crucial to interpreting LLM outputs with a critical lens.

Unlocking the potential of LLMs: the art of prompt engineering

Getting the best out of LLMs requires mastering the art of prompt engineering. The clearer and more specific your prompt is, the more accurate the outcome. Here are some key tips for crafting impactful prompts:

- **Be crystal clear:** Avoid ambiguity and unnecessary fluff. Use complete sentences and precise language to convey your intent.

- **Know your goal:** Is your LLM assisting with research, generating patient summaries, or composing educational materials? Tailor your prompt accordingly.
- **Provide context:** LLMs thrive on background information. Offer relevant facts, definitions, or even previous dialogues to set the stage.
- **Target with keywords:** Sprinkle in relevant keywords like 'urology', 'diagnosis' or 'treatment' to guide the LLM towards your desired area.
- **Break it down:** Complex problems benefit from smaller, more specific prompts. Instead of asking for a 'surgical report' focus on specific aspects like 'preoperative findings' or 'postoperative instructions'.
- **Embrace examples:** If you have a clear vision, provide examples to illustrate your desired outcome. This is especially helpful for creative tasks like writing

DIGITAL REVIEW

medical summaries or patient education materials.

- **Proofread and iterate:** Don't underestimate the power of a second look! Typos and grammatical errors can throw the LLM off track. Revise your prompt as needed, remembering you can always refine it based on the initial response.

LLMs in urology: a glimpse of the future

The potential applications of LLMs in urology are vast. Imagine LLMs analysing medical images to detect early signs of prostate cancer, summarising complex research papers for busy clinicians, or even creating personalised patient education materials in multiple languages. Google's Med-Palm 2, a medical-specific LLM is being tested in improving clinical trials, reducing administrative burden by extracting information from doctor-patient conversations, and generating medical notes [1]. Med-Palm 2 even managed to achieve an impressive score of 85% on the United States Medical License Exam (USMLE) [2]. This glimpse into the future underscores the immense potential of LLMs to revolutionise urology by:

- **Enhancing diagnostics:** LLMs can analyse medical records, scans and images to support faster and more

accurate diagnoses, potentially aiding in early disease detection.

- **Streamlining research:** By digesting vast amounts of research data and identifying relevant trends, LLMs has the potential to accelerate research efforts and inform clinical decision-making.
- **Optimising patient care:** LLMs can personalise patient education materials, answer questions, and even offer emotional support, contributing to improved patient engagement and adherence to treatment plans.
- **Empowering clinicians:** LLMs can alleviate administrative burdens by extracting key information from patient consultations and generating medical reports, freeing up clinicians' time for patient care.

The road ahead: collaboration and innovation

The LLM revolution in healthcare is still in its early stages. Collaboration between clinicians, researchers, and AI developers is crucial to ensure responsible development and ethical implementation of these powerful tools. As we refine LLMs and address their limitations, the future promises to be one of enhanced accuracy, efficiency and, ultimately, improved patient care.

Remember, the real magic lies not just in the LLM itself, but in its synergy with human expertise. So, embrace the opportunity, delve into the world of LLMs, and explore their potential to transform the future of healthcare!

References

1. <https://cloud.google.com/blog/topics/healthcare-life-sciences/introducing-medlm-for-the-healthcare-industry> [Accessed 22 January 2024].
2. Singhal K, Azizi S, Tu T, et al. Large language models encode clinical knowledge. *Nature* 2023;**620**(7972):172–80.

Further reading

<https://medium.com/data-science-at-microsoft/how-large-language-models-work-91c362f5b78f>

<https://copilot.microsoft.com>

<https://gemini.google.com>

<https://chat.openai.com>

AUTHOR AND SECTION EDITOR



Ivo Dukic,

Consultant Urological Surgeon, University Hospitals Birmingham.