



Kenapa ChatGPT Bisa Memberikan Jawaban Palsu atau Tidak Faktual?

Swipe for more





MATTHEWS WONG



Dalam sebuah pidato baru-baru ini, Prof. Stella Christie, yang saat ini menjabat sebagai Wakil Menteri Pendidikan Tinggi, Sains, dan Teknologi, menyoroti risiko informasi palsu dari ChatGPT. Ia membagikan contoh di mana ChatGPT mengutip sebuah makalah penelitian yang sebenarnya tidak ada. **Jadi, mari kita telusuri—mengapa hal ini bisa terjadi?**

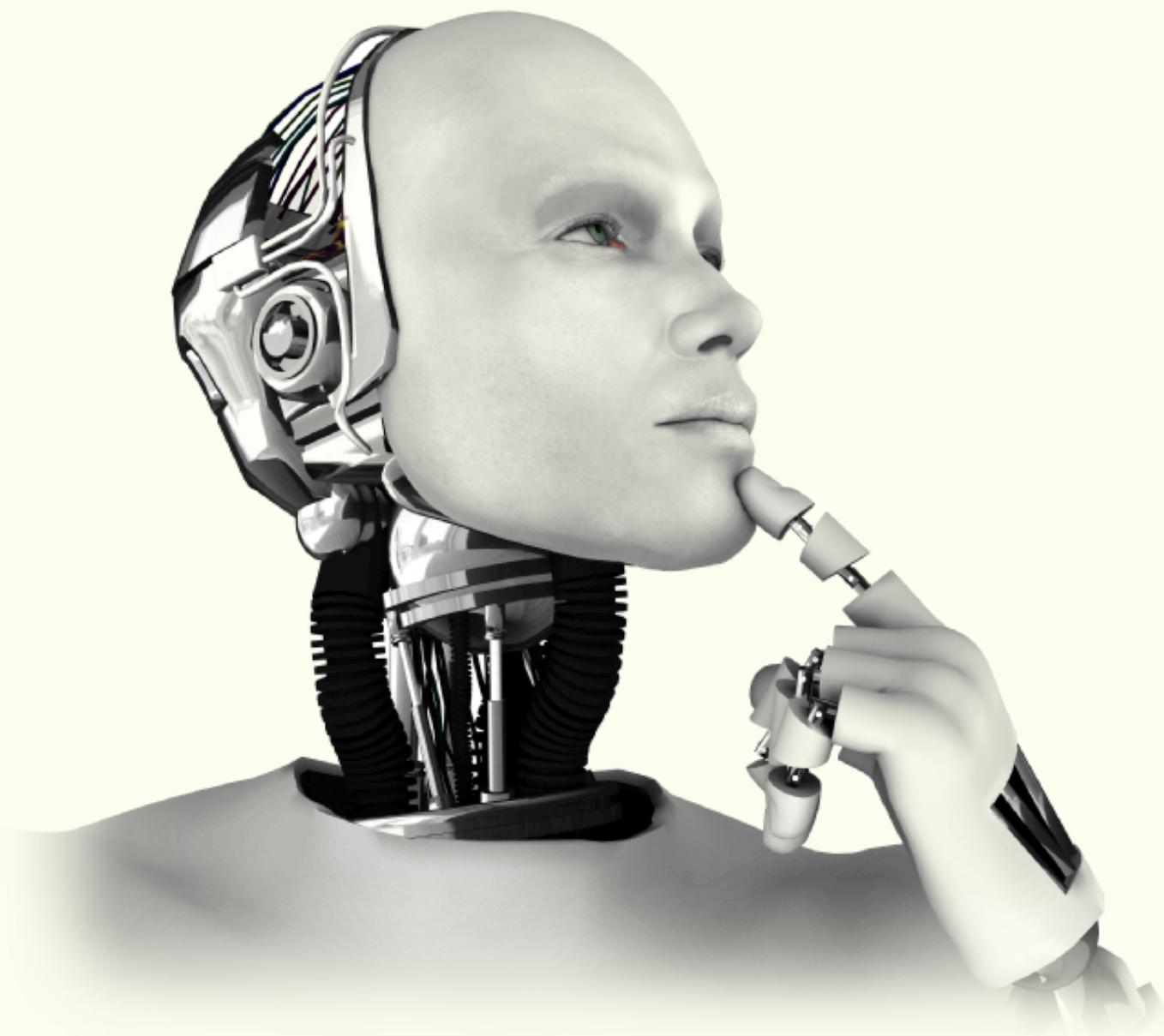




MATTHEWS WONG

Itu disebabkan oleh:

LLM HALLUCINATION



Swipe for more



Apa itu LLM Hallucination ?

- LLM Hallucination terjadi saat model menghasilkan teks yang tampak relevan namun tidak akurat karena tidak memiliki referensi data nyata.
- Hallusinasi muncul ketika model memprediksi kata atau frasa berdasarkan pola pelatihan statistik, bukan dari pemahaman kontekstual yang benar.
- Bias ini sering terjadi dalam situasi di mana model tidak memiliki data pelatihan yang cukup spesifik atau menghadapi pertanyaan ambigu di luar cakupan datasetnya.





Penyebab Umum *LLM Hallucination*

- Model AI hanya mengandalkan data yang ada, sehingga bisa menghasilkan jawaban tidak akurat.
- Model AI memprediksi teks tanpa pemahaman mendalam, dan sering mengabaikan konteks.
- Tidak Ada Verifikasi Fakta: Model tidak memiliki cara untuk memeriksa kebenaran jawaban yang dihasilkan.



Cara menghindari LLM Hallucination ketika menggunakan chatbot AI

- Gunakan Teknik *Chain of Thought* Prompting: Instruksikan AI untuk merinci langkah berpikirnya dalam menjawab, membantu menghasilkan respons yang lebih logis dan terstruktur.
- Gunakan Konteks Tambahan pada prompting chatbot : Tanyakan kembali atau tambahkan konteks spesifik untuk memandu AI, mengurangi kemungkinan jawaban yang tidak relevan.



Hal yang bisa dipelajari adalah :

- Jangan percaya sepenuhnya kepada AI dan lakukan cek ulang pada sumber yang terpercaya, dan ketika ada paper yang dikasih bisa cek di Google Scholar atau semacamnya.
- Gunakan chatbot untuk menjadi teman berpikir, bukan untuk copy paste.
- Harus lebih kritis ketika mendapatkan jawaban dari AI karena itu belum tentu benar.





Referensi :

- E. Samson, "What Are LLM Hallucinations – Examples and Causes," The Gleen Blog: Where Generative AI & CS/CX Intersect, Dec. 2023. <https://gleen.ai/blog/llm-hallucination/>
- M. Bilan, "Hallucinations in LLMs: What You Need to Know Before Integration," Master of Code Global, Jul. 14, 2023. <https://masterofcode.com/blog/hallucinations-in-llms-what-you-need-to-know-before-integration>

