Easily double-check responses and build on shared conversations

Google Bard Updates

What is Hallucination

Hallucination in the context of LLMs refers to the generation of text that is erroneous, nonsensical, or detached from reality.

LLM Hallucination Rate

Maintenance and reliability knowledge exam	
ChatGPT	5%
Bard	4%

PHM knowledge exam	
ChatGPT	10%
Bard	29%

Types of Hallucinations

Causes and Types of LLMs Hallucination

Causes of LLMs Hallucination

- Source-Reference Divergence
- Exploitation through Jailbreak Prompts
- Reliance on Incomplete or Contradictory Datasets
- Overfitting and Lack of Novelty
- Guesswork from Vague or Insufficiently Detailed Prompts

Types of LLMs Hallucination

- Sentence Contradiction
- Prompt Contradiction
- Factual Contradiction
- Nonsensical Output
- Irrelevant or Random Hallucinations

Consequences of LLM Hallucination: Addressing Ethical Concerns

Privacy Issues

6x increase in toxicity: ChatGPT persona assignment perpetuates harmful stereotypes

Misinformation and Disinformation

74% of IT decision-makers concerned about cybersecurity risks with GPT-4

Discriminating and Toxic Content

49% of individuals foresee GPT-4 as a tool for spreading misinformation

Easily double-check responses and build on shared conversations