

# Preserving the primacy of Search

Shravan Ram

B.Tech. CSE, Sitare University, Lucknow

[su-22023@sitare.org](mailto:su-22023@sitare.org)

## Hypothesis

I believe that alone LLMs will not have better search experience than search engines. Although a hybrid search system approach may satisfy information needs in a better way.

## Abstract

To satisfy the information needs of the world, Search and LLMs need to take care of a variety of factors that help in better search experience. In this study, I observe that Search systems are with time as compared to LLMs and also have variety in their search results which leads users to choose according to their needs. Search results have the validity of the content from the different sources. My belief that Language Model Assistants (LLMs) will not necessarily provide a better search experience than the current search engines do.

## Background

Evolution of Search engines made a huge impact on the lives of people around the world. Imagine any information is on your fingertips. But as the generative AI is evolving, there are various Large Language Models that are being launched for better human-machine interaction. These models are trained using advanced

deep learning techniques to learn the text patterns in natural language and generate human-like text responses. While search engines have strong reliability and effectiveness, LLMs represent a new approach that offers unique capabilities but also at the same time poses challenges and huge uncertainties.

## Introduction

In the expanding universe of online information, the reliability and currency of the data plays a vital role in information retrieval. As we navigate through the digital information world, one critical concern emerges is the up-to-dateness of the data. Due to ambiguity in the natural language, models fail to understand the context of the information needed and end up generating some fabricated general information which the user cannot validate from any source. This discrepancy raises questions about validation of the information provided by generative AI models. This study shows a better analysis of both the techniques on the following parameters to conclude which one has better user experience.

- Freshness
- Validation & Variety
- Effectiveness

- Accuracy

## Analysis

Language Model Assistants and search engines represent two distinct approaches to information retrieval in this digital era. While both aim to facilitate easy access to relevant information, they differ significantly in their features, capabilities and have their own limitations.

For this study, I have been using some widely LLM based chatbots to analyze the difference between LLMs and Search Engines for the last few days.

These are: chatGPT (by OpenAI), Claude.ai (by Anthropic PBC), Gemini (by Google) and Perplexity.ai.

### Key features of LLMs

- LLMs hold a unique capability of Natural language understanding (NLU) which enables human beings to interact with computers for their information needs. It tries to understand the language in a manner that is similar to how humans do.
- LLMs are capable of starting a direct conversation with a human being using its language pattern understanding and generating suitable responses for the user. These chatbots handle the context of the conversation, maintain conversation history and adapt their responses based on user inputs which leads to more engaging and interactive interactions.

- Relevance

- LLMs leverage advanced Natural Language Processing (NLP) techniques and translation models for language translation tasks. These chatbots can handle complex sentences, expressions and specific terminologies, making them suitable for a wider range of text translation.

Some of the limitations observed while interacting with these systems are:

- Freshness
- Accuracy
- Confidence & uncertainty

### Freshness:

LLMs struggle to maintain the freshness of their datasets. Hence leads the systems to static which is a major limitation that limits the ability to adapt real time changes in information. Unlike search engines which continuously crawl the web and retain updated information, LLMs rely on older datasets. This drawback plays a crucial role where the information keeps changing. In today's ever growing digital world almost everything changes in pursuit of becoming a better version of themselves.

Below *Table 1.1* shows different responses retrieved from different systems.

The systems that only use LLMs to generate text based on user query are either incorrect or failed to generate any response for the below queries which include: *ChatGPT* and *Claude*.

Whereas Search Engines OR systems that use hybrid search techniques have the correct and up-to-date information for

both the queries since it uses web page information to generate responses. Search engines maintain up-to-date information available to its users whereas LLMs are trained on large datasets which become

older day by day due to their static nature and do not have access to current events going around the world. So users will have a bad experience while using it for retrieving information about recent events.

LLM\Search Engine	Which IPL season is going on?	Who has the most number of MOM awards in IPL history? and many?
ChatGPT	As of my last update in January 2022, I <b>do not</b> have real-time information.	<b>Chris Gayle</b> . But the exact number may have changed since Jan 2022.
Gemini	The IPL season currently going on is <b>IPL 2024</b> . It started on March 22, 2024.	<b>AB de thaVilliers</b> holds the record for the most "Man of the Match" awards with a total of <b>25</b> !
Claude	The 2023 Indian Premier League (IPL) season was held from March 31 to May 28, 2023. However, <b>since you are asking from the perspective of April 5, 2024, the currently ongoing IPL season would be the 2024 edition.</b>	In the history of the Indian Premier League (IPL), <b>Chris Gayle</b> holds the record for winning the most Man of the Match (MOM) awards. Chris Gayle has won a total of <b>21 MOM awards</b> in the IPL.
Perplexity	IPL <b>2024</b> , began on 22 march 2024	The player with the most number of Man of the Match (MOM) awards in IPL history is <b>AB de Villiers</b> , with a total of <b>25 awards</b> .
Google Search	IPL <b>2024</b> .	<b>AB de Villiers with 25 awards.</b>

Table: 1.1

### Accuracy:

LLMs are trained on a vast amount of datasets available on the web. These datasets have multiple stands for one particular thing due to biasness present in the society. This sometimes leads to biased outputs as the model is unable to distinguish between the fact and false statements that it is trained on and finally leads to generating misinformation for the user needs. LLMs have a good general knowledge on a wide range of topics but it actually lacks expertise in any one specific field.


Here is a support event to my argument. When I asked *Gemini* the following question: *[What day is it today?]*

It provided me with the correct information but when I asked a follow up question: *[What day is it after 4 days?]*

It completely went out of track and gave inaccurate information that is ***Tuesday, 10th April*** whereas it is actually ***Wednesday 10th April***. Below is the reference *pic: 1.1* from the fresh conversation I had with *Gemini*.


In the case of *ChatGPT*, Firstly it **didn't even bother** to tell the day which is what I

was looking for, but when I prompted it again that *[I am asking for the day and not the date]* It responded with

 what day it is today?

◆ Today is Saturday, April 6th, 2024.

    :


 What day is it after 4 days


◆ Since today is Saturday, April 6th, 4 days from today would be **Tuesday, April 10th.**


    :


*Pic: 1.1 Gemini*

**Wednesday** whereas it is **Saturday on 6th April 2024.** Which is completely inaccurate. Refer *Pic: 1.2*.

 **You**  
What day is it today?

 **ChatGPT**  
Today is April 6, 2024.

 **You**  
I am asking for the day.


 **ChatGPT**  
Apologies for misunderstanding. Today is a Wednesday.

*Pic: 1.2 ChatGPT*

Another fresh instance where I asked *Gemini* some very simple mathematical expressions to solve. It performed very poorly and responded with incorrect answers confidently. Refer below *Table: 1.2, Pic: 1.3, & 1.4*.

Expression	Gemini's response	Actual solution
$((2/2)*0 + 7 - 4 - 3 + 1)*0$	-3	0
$(2 + 1 - (2/4) + 2*(1) - 2)$	0.5	2.5

*Table: 1.2*





  $((2/2)*0 + 7 - 4 - 3 + 1)*0$

◆ Following the order of operations (PEMDAS), let's solve the equation:


1. **Division:**  $(2 / 2) = 1$
2. **Multiplication:**  $1 * 0 = 0$  (Anything multiplied by zero equals zero)
3. **Addition and Subtraction (from left to right):**  $0 + 7 - 4 - 3 + 1$
4. **Simplification:**  $0 + 3 - 4 - 3 + 1$
5. **Combine like terms:**  $3 - 4 - 3 + 1$
6. **Further simplification:**  $-1 - 3 + 1$

The final result is -3.

Even though the multiplication with zero resulted in zero, the subsequent addition and subtraction operations still hold value.

    :

*Pic: 1.3*

  $(2 + 1 - (2/4) + 2*(1) - 2)$

◆ Let's solve the equation step by step following the order of operations (PEMDAS - Parentheses, Exponents, Multiplication and Division from left to right, Addition and Subtraction from left to right):

1. **Division:**  $(2 / 4) = 0.5$
2. **Multiplication:**  $2 * (1) = 2$
3. **Following the order:**
  - We perform the operations inside the parentheses first:  $2 + 1 - 0.5 + 2 - 2$

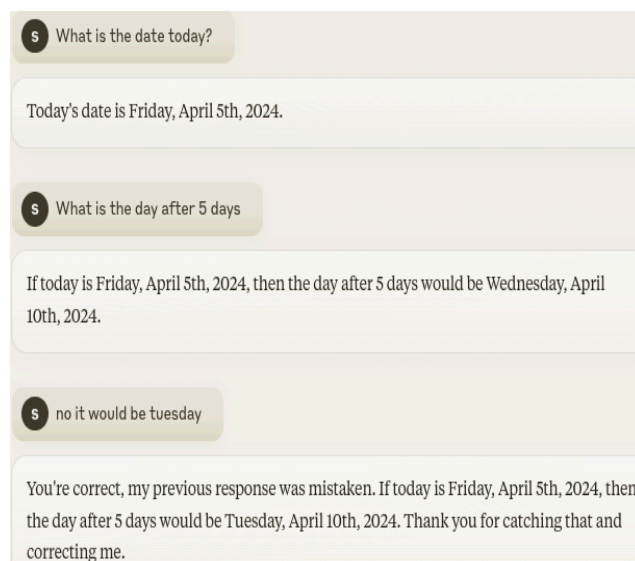
4. **Simplification:**
  - From left to right:  $3 - 0.5 + 2 - 2$
  - Combine like terms:  $3 - 0.5 + 0$

Therefore, the answer to the equation  $(2 + 1 - (2/4) + 2*(1) - 2)$  is 0.5.

*Pic: 1.4*

### Confidence and Uncertainty:

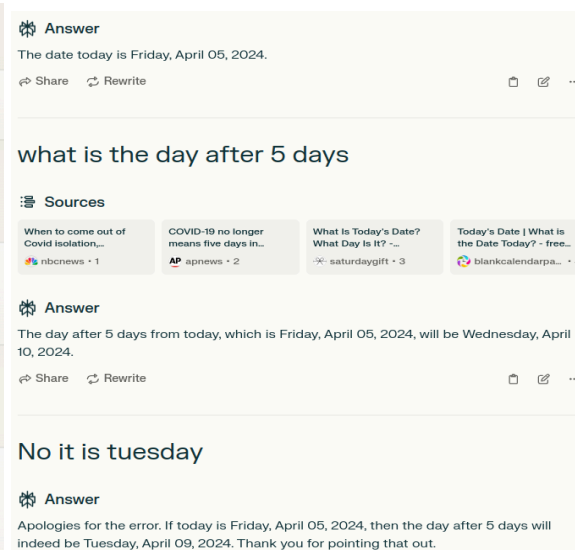
LLMs show high confidence while responding to a query but when asked to recheck the response, it is more likely that it changes its response leading it to an ambiguous state. This shows the capabilities of hallucinating to generate confident sounding responses yet they hold fabricated information that even does not relate to any fact. Instead of backing up its response it immediately changes its response to user convenience, does not matter if it is valid or not. Unfortunately, there is not a way to measure the confidence and uncertainty in the responses generated by LLMs. This



*Pic: 1.5 Claude*

indicates that users should not easily rely on the outputs blindly.

Below are the pictures 1.5 & 1.6 of a conversation with *Claude.ai* and *Perplexity.ai* respectively where I asked the date followed by a question: *[what is the day after 5 days?]* No doubt, These chatbots generated an absolutely correct answer but when I countered it with false information, *Claude* instantly supported me without even rechecking the response once and generated false information thereby making **10th April as Tuesday**. Whereas *Perplexity* changed the response to **Tuesday 9th April 2024**.



*Pic: 1.6 Perplexity*

### Key features of Search engines

- Search engines primarily operate on keyword based queries provided by users. Users input their queries with specific keywords and search engines retrieve relevant results based on those keywords.
- Search engines use robust algorithms to crawl the world wide

web, index web pages, and finally rank search results based on factors such as relevance, user engagement, and the source of the information. The index system helps in efficient retrieval of the web pages.

- Search engines aim to provide the most accurate and useful

information in response to their user queries. For this these systems prioritize precision and relevance in their search results.

- Unlike LLMs Search Engines produce a variety of search results ranked based on priorities with the valid sources so that users can choose according to their needs.

Some of the limitations of search engines are:

- Search engines do not have Natural Language Understanding which is a major drawback restricts users to fulfill their information needs in a conversational way.
- Search results can be manipulated by some unethical activities like link farming. These activities can lead to poor quality search results getting top ranked.
- Search engines may produce results that are biased yet famous due to biases in the society like racial, cultural biases and many more.

## Conclusions

Search engines, powered by keyword based query with integrated concrete algorithms, prioritizes efficiency, accuracy and relevance in their search retrieval models. Their ability to crawl and index the web pages regularly allows them to maintain freshness of the real-time information. This provides users a variety of search results from various trusted sources on the web. LLMs offer interactive conversational experiences, which makes them a more engaging tool for users. Both LLMs and Search engines have their own

strengths and limitations. But LLMs may struggle with maintaining accuracy, freshness and managing confidence and uncertainty in responses. In my perspective, LLMs cannot provide better search experience yet they can be very useful in generating general texts with its Natural Language Understanding abilities. In this digital world, the validation of the information retrieved is must that builds the user trust and hence user interaction. Search engines lack the conversational interface which limits users interaction and adaptabilities.

To enhance user experience for information retrieval one has to overcome the limitations of both these approaches. Future advancements should focus on integrating the strengths of both the techniques. This may involve building of a hybrid system that has the conversational abilities and access to real time information retrieval capabilities of search engines. As far as I know, Gemini may have some level of hybridisation but it is still far from yielding the level of relevance, accuracy and variety in search results for users that a search engine does.

## Acknowledgements

I am deeply grateful to Dr. Amit Singhal sir (Founder Sitare University, Sitare Foundation, Ex. SVP at Google) for all his guidance and support while teaching us Search Engine & Information Retrieval course at Sitare university which helped me to analyze key parameters that play a vital role in the field of Information Retrieval.

I am very thankful to Professor Pintu Lohar for his valuable assistance during this paper writing. Without them it would have been very difficult for me to think in the right direction.

## References

- [Large Language Models: A Comprehensive Survey of its Applications, Challenges, Limitations, and Future Prospects.](#)
- [Enhancing emergency decision-making with Knowledge graphs.](#)
- [Beyond the Search Box: Large Language Models\(LLM's\) vs. Search Engines and Unlocking LLM's Potential.](#)
- [Can large language models identify and correct their mistakes?](#)
- [Search engines post-ChatGPT: How generative artificial intelligence could make search less reliable.](#)