Search

[Write](https://medium.com/new-story?source=---two_column_layout_nav----------------------------------)

Rishabh Sahai

Get unlimited access to the best of Medium for less than $1/week.

[Become a member](https://medium.com/plans?source=upgrade_membership---post_top_nav_upsell----------------------------------)

Top highlight

**Knowledge Graphs & LLMs: Fine-Tuning Vs. Retrieval-Augmented Generation**

What are the limitations of LLMs, and how to overcome them

[[](https://medium.com/@bratanic-tomaz?source=post_page-----30e875d63a35--------------------------------)](https://medium.com/@bratanic-tomaz?source=post_page-----30e875d63a35--------------------------------)

[[Neo4j Developer Blog](https://medium.com/neo4j?source=post_page-----30e875d63a35--------------------------------)](https://medium.com/neo4j?source=post_page-----30e875d63a35--------------------------------)

[Tomaz Bratanic](https://medium.com/@bratanic-tomaz?source=post_page-----30e875d63a35--------------------------------)

·

Follow

Published in

[Neo4j Developer Blog](https://medium.com/neo4j?source=post_page-----30e875d63a35--------------------------------)

·

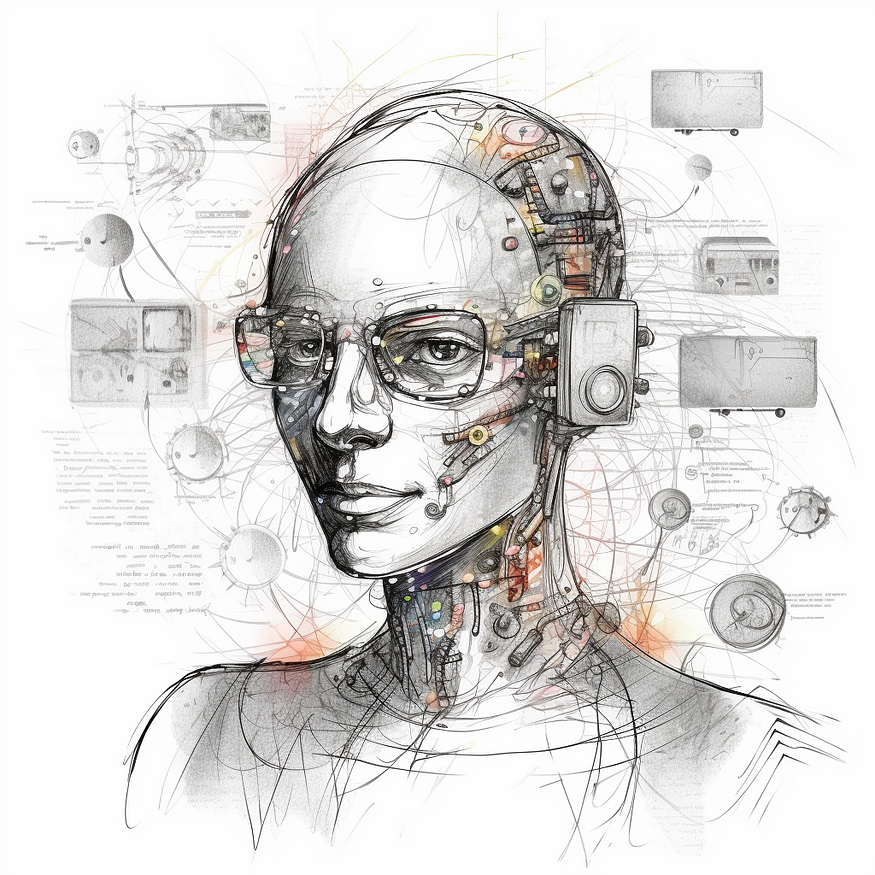
12 min read

·

Jun 6, 2023

719

7



Midjourney’s idea of a knowledge graph chatbot.

*This is the second blog post of Neo4j’s NaLLM project. We started this project to explore, develop, and showcase*[*practical uses of these LLMs in conjunction with Neo4j*](https://neo4j.com/generativeai/?utm_campaign=gen-ai&utm_content=medium&utm_source=blog)*. As part of this project, we will construct and publicly display demonstrations in a*[*GitHub repository*](https://github.com/neo4j/NaLLM)*, providing an open space for our community to observe, learn, and contribute. Additionally, we are writing about our findings in a blog post. You can find the first and third blog posts here:*

* [Harnessing LLMs With Neo4j](https://medium.com/neo4j/harnessing-large-language-models-with-neo4j-306ccbdd2867)
* [Multi-Hop Question Answering](https://medium.com/neo4j/knowledge-graphs-llms-multi-hop-question-answering-322113f53f51)

**[Harnessing Large Language Models with Neo4j](https://medium.com/neo4j/harnessing-large-language-models-with-neo4j-306ccbdd2867?source=post_page-----30e875d63a35--------------------------------" \t "_blank)**

[Episode 1 — Exploring Real-World Use Cases](https://medium.com/neo4j/harnessing-large-language-models-with-neo4j-306ccbdd2867?source=post_page-----30e875d63a35--------------------------------" \t "_blank)

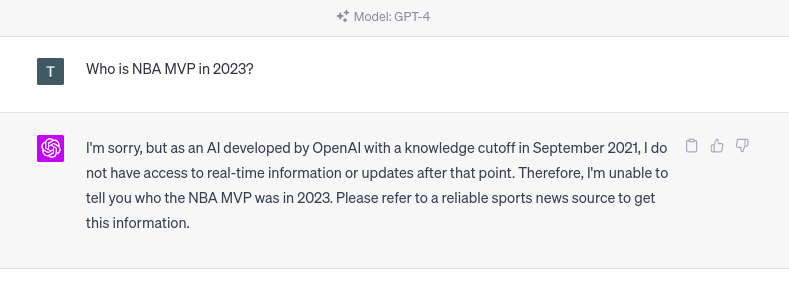
[medium.com](https://medium.com/neo4j/harnessing-large-language-models-with-neo4j-306ccbdd2867?source=post_page-----30e875d63a35--------------------------------" \t "_blank)

The first wave of hype for Large Language Models (LLMs) came from ChatGPT and similar web-based chatbots, where the models are so good at understanding and generating text that it shocked people, myself included.

Many of us logged in and tested its ability to write haikus, motivational letters, or email responses. What became quickly apparent is that LLMs are not only good at generating creative context but also at solving typical natural language processing and other tasks.

Shortly after the LLM hype started, people started considering integrating it into their applications. Unfortunately, if you simply develop a wrapper around an LLM API, there is a high chance your application will not be successful as it doesn’t provide additional value.

One major problem of LLMs is the so-called **knowledge cutoff**. The knowledge cutoff term indicates that LLMs are ***unaware of any events that happened after their training***. For example, if you ask ChatGPT about an event in 2023, you will get the following response.



ChatGPT’s knowledge cutoff date. Image by author.

The same problem will occur if you ask an LLM about ***any event not present*** in its training dataset. While the knowledge cutoff date is relevant for any publicly available information, the LLM doesn’t have any knowledge about ***private or confidential information*** that might be available even before the knowledge cutoff date.

For example, most companies have some confidential information that they don’t share publicly but might be interested in having a custom LLM that could answer those questions. On the other hand, a lot of the publicly available information that the LLM is aware of might be already outdated.

*Therefore, updating and expanding the knowledge of an LLM is highly relevant today.*

Another problem with LLMs is that they are trained to produce realistic-sounding text, which ***might not be accurate***. Some invalid information is more challenging to spot than others. Especially for missing data, it is very probable that the LLM will make up an answer that sounds convincing but is nonetheless wrong instead of admitting that it lacks the base facts in its training.

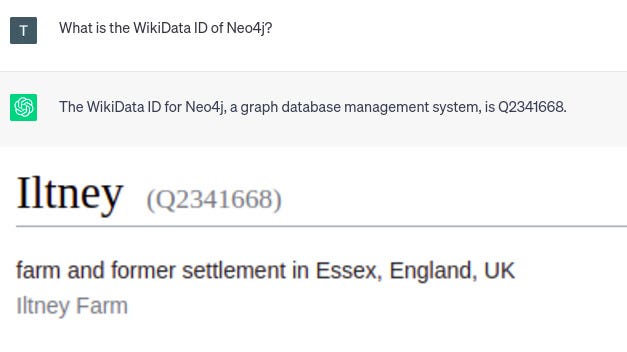
For example, research or court citations might be easier to verify. A week ago, a [lawyer got in trouble for blindly believing the court citations ChatGPT produced](https://edition.cnn.com/2023/05/27/business/chat-gpt-avianca-mata-lawyers/index.html).

**[Lawyer apologizes for fake court citations from ChatGPT | CNN Business](https://edition.cnn.com/2023/05/27/business/chat-gpt-avianca-mata-lawyers/index.html?source=post_page-----30e875d63a35--------------------------------" \t "_blank)**

[The meteoric rise of ChatGPT is shaking up multiple industries - including law. A lawyer for a man suing Avianca…](https://edition.cnn.com/2023/05/27/business/chat-gpt-avianca-mata-lawyers/index.html?source=post_page-----30e875d63a35--------------------------------" \t "_blank)

[edition.cnn.com](https://edition.cnn.com/2023/05/27/business/chat-gpt-avianca-mata-lawyers/index.html?source=post_page-----30e875d63a35--------------------------------" \t "_blank)

I have also noticed that LLMs will consistently ***produce assertive, yet false information about any sort of IDs*** like the WikiData or other identification numbers.



ChatGPT’s hallucinations. Image by author.

Since the response by ChatGPT is assertive, you might expect it to be accurate. However, the given WikiData id points to a farm in England. Therefore, you have to be very careful not to blindly believe everything that LLMs produce. Verifying answers or producing more accurate results from LLMs is another big problem that needs to be solved.

Of course, LLMs have other problems, like bias, prompt injection, and others. However, we will not talk about them here. Instead, in this blog post, we will present and focus on the ***concepts of fine-tuning and retrieval-augmented LLMs*** and evaluate their pros and cons.

**Supervised Fine-Tuning of an LLM**

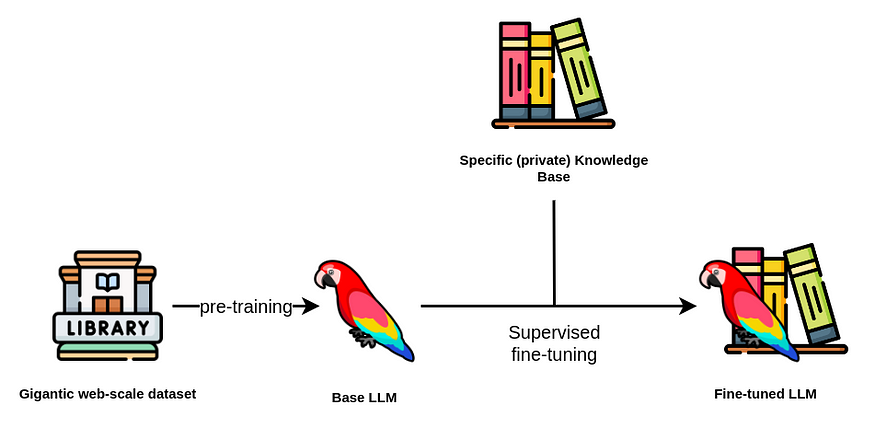
Explaining how LLMs are trained is beyond the scope of this blog post. Instead, you can watch this [incredible video by Andrej Karpathy to catch up on LLMs](https://www.youtube.com/watch?v=bZQun8Y4L2A) and learn about the different phases of LLM training.

*By fine-tuning an LLM, we refer to the***supervised training phase***, during which you provide additional question-answer pairs to optimize the performance of the Large Language Model (LLM).*

Additionally, we have identified two different use cases for fine-tuning an LLM.

**One use case** is fine-tuning a model to update and ***expand its internal knowledge***.  
In contrast, the **other use case** is focused on fine-tuning a model ***for a specific*** task like text summarization or translating natural language to database queries.

First, we will talk about the first use case, where we use fine-tuning techniques to update and expand the internal knowledge of an LLM.



Supervised fine-tuning flow. Image by author. Icons from [Flaticons](https://www.flaticon.com/" \t "_blank).

Usually, you want to avoid pre-training an LLM as the **cost** can be upwards of hundreds of thousands and even millions of dollars. A base LLM is pre-trained using a gigantic corpus of text corpus, frequently in the billions or even trillions of tokens.

While the **number of parameters** of an LLM is vital, it is not the only parameter you should consider when selecting a base LLM. Besides the **license**, you should also consider the **bias** and **toxicity** of the pre-training dataset and the base LLM.

After you have selected the base LLM, you can start the next step of fine-tuning it. The fine-tuning step is relatively cheap regarding computation cost due to available techniques like the [LoRa](https://huggingface.co/blog/lora) and [QLoRA](https://arxiv.org/abs/2305.14314" \t "_blank).

**[Using LoRA for Efficient Stable Diffusion Fine-Tuning](https://huggingface.co/blog/lora?source=post_page-----30e875d63a35--------------------------------" \t "_blank)**

[LoRA: Low-Rank Adaptation of Large Language Models is a novel technique introduced by Microsoft researchers to deal…](https://huggingface.co/blog/lora?source=post_page-----30e875d63a35--------------------------------" \t "_blank)

[huggingface.co](https://huggingface.co/blog/lora?source=post_page-----30e875d63a35--------------------------------" \t "_blank)

However, ***constructing a training dataset*** is more complex and can get expensive. If you can not afford a dedicated team of annotators, it seems that the trend is to ***use an LLM to construct a training dataset*** to fine-tune your desired LLM (this is really meta).

For example,[Stanford’s Alpaca training dataset was created using OpenAI’s LLMs](https://crfm.stanford.edu/2023/03/13/alpaca.html). The cost to produce 52 thousand training instructions was about 500 dollars, which is relatively cheap.

**[Stanford CRFM](https://crfm.stanford.edu/2023/03/13/alpaca.html?source=post_page-----30e875d63a35--------------------------------" \t "_blank)**

[We introduce Alpaca 7B, a model fine-tuned from the LLaMA 7B model on 52K instruction-following demonstrations. On our…](https://crfm.stanford.edu/2023/03/13/alpaca.html?source=post_page-----30e875d63a35--------------------------------" \t "_blank)

[crfm.stanford.edu](https://crfm.stanford.edu/2023/03/13/alpaca.html?source=post_page-----30e875d63a35--------------------------------" \t "_blank)

On the other hand, the [Vicuna model was fine-tuned by using the ChatGPT conversations users posted on ShareGPT.com](https://lmsys.org/blog/2023-03-30-vicuna/).

**[Vicuna: An Open-Source Chatbot Impressing GPT-4 with 90%\* ChatGPT Quality | LMSYS Org](https://lmsys.org/blog/2023-03-30-vicuna/?source=post_page-----30e875d63a35--------------------------------" \t "_blank)**

[by: The Vicuna Team, Mar 30, 2023 We introduce Vicuna-13B, an open-source chatbot trained by fine-tuning LLaMA on…](https://lmsys.org/blog/2023-03-30-vicuna/?source=post_page-----30e875d63a35--------------------------------" \t "_blank)

[lmsys.org](https://lmsys.org/blog/2023-03-30-vicuna/?source=post_page-----30e875d63a35--------------------------------" \t "_blank)

There is also a relatively [fresh project by H2O called WizardLM](https://github.com/h2oai/h2o-wizardlm), which is designed to turn documents into question-answer pairs that can be used to fine-tune an LLM.

We haven’t found any recent articles describing how to use a knowledge graph to prepare good question-answer pairs that can be used to fine-tune an LLM.

This is an area that we plan to explore during the NaLLM project. We have some ideas for utilizing LLMs to construct question-answer pairs from a knowledge graph context.

**However, there are a lot of unknowns at the moment.**  
For example, can you provide *two different answers* to the same question, and the LLM then somehow combines them in its internal knowledge store?

Another consideration is that some information in a knowledge graph is not relevant without considering its relationships. Therefore, do we have to pre-define relevant queries, or is there a more generic way to go about it? Or can we use the node-relationship-node patterns representing subject-predicate-object expressions to generate relevant pairs?

These are some of the questions we aim to answer in upcoming blog posts.

Imagine that you somehow managed to produce a training dataset containing question-answer pairs based on the information stored in your knowledge graph. As a result, the LLM now includes updated knowledge.

However, fine-tuning the model didn’t solve the knowledge cutoffs problem since it only pushed the knowledge cutoff to a later date.

Therefore, we recommend updating the internal knowledge of an LLM through fine-tuning techniques *only for slowly changing*or updating data. For example, you could use a fine-tuned model to provide tourist information.

However, you would run into troubles the second you would want to include special time-dependent (real-time) or personalized promotions in the responses. Similarly, fine-tuned models are not ideal for analytical workflows where you would ask how many new customers the company gained over the last week.

At the moment, *fine-tuning approaches can help mitigate hallucinations* but cannot completely eliminate them. One problem is that the LLMs *do not cite their sources* when providing answers. Therefore, you have no idea if the answer came from pre-training data, fine-tuning dataset, or was made up by the LLM. Additionally, there might be another possible falsehood source if you use an LLM to create the fine-tuning dataset.

Lastly, a fine-tuned model cannot automatically provide different responses *depending on the user* making the questions. Likewise, there is no concept of access restrictions, meaning that anybody interacting with the LLM has access to all of its information.

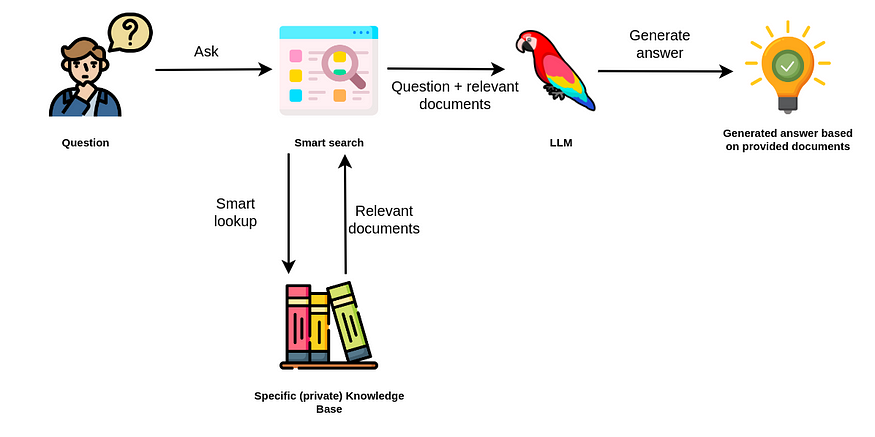
**Retrieval-Augmented Generation**

Large language models perform remarkably well in natural language applications like

* Text summarization,
* Extracting relevant information,
* Disambiguation of entities
* Translating from one language to another, or even
* Converting natural language into database queries or scripting code.

Moreover, previously NLP models were most often domain and task-specific, meaning that you would most likely need to train a custom natural language model depending on your use case and domain. However, thanks to the generalization capabilities of LLMs, a single model can be applied to solve various collections of tasks.

We have observed quite a strong trend in using retrieval-augmented LLMs, where instead of using LLMs to access its internal knowledge, you use the LLM ***as a natural language interface*** to your company’s or private information.



Retrieval-augmented generation. Image by author. Icons from [Flaticons](https://www.flaticon.com/" \t "_blank).

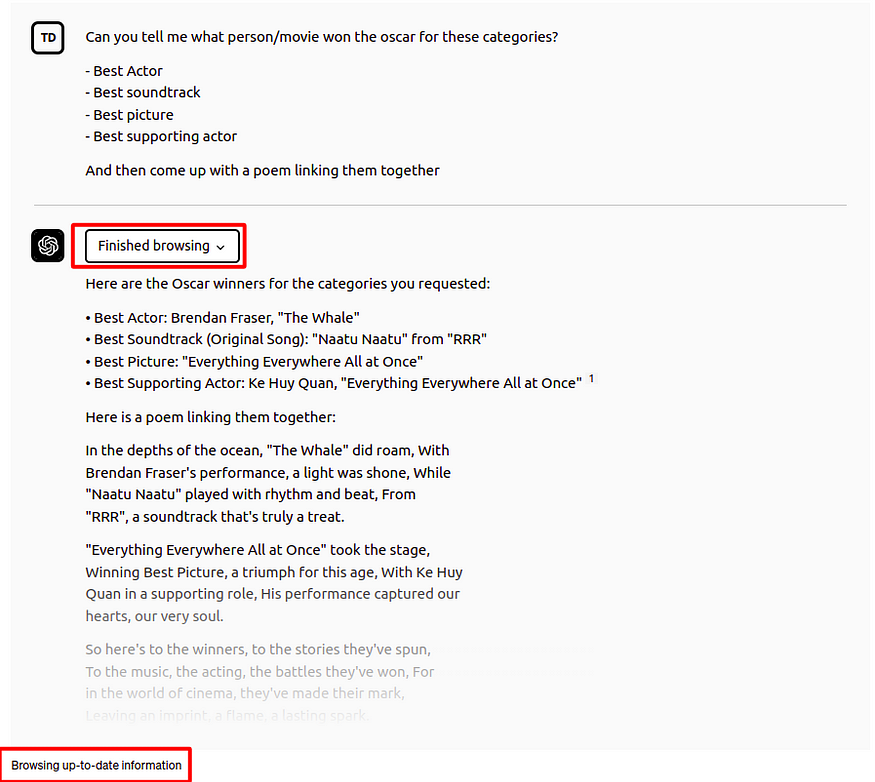
The retrieval augmented approach uses the LLM to generate an answer based on the additionally provided relevant documents from your data source.

Therefore, you don’t rely on internal knowledge of the LLM to produce answers. Instead, the LLM is used only for extracting relevant information from documents you passed in and summarizing it.

**[ChatGPT plugins](https://openai.com/blog/chatgpt-plugins?source=post_page-----30e875d63a35--------------------------------" \t "_blank)**

[openai.com](https://openai.com/blog/chatgpt-plugins?source=post_page-----30e875d63a35--------------------------------" \t "_blank)

For example, the [ChatGPT plugins](https://openai.com/blog/chatgpt-plugins) can be thought of as a retrieval-augmented approach to LLM applications. The ChatGPT interface with a browsing plugin enabled allows the LLM to search the internet to access up-to-date information and use it to construct the final answer.



ChatGPT with browsing plugin. Image by author.

In this example, ChatGPT was able to answer who won the Oscar for various categories in 2023. But, remember, the cutoff knowledge date for ChatGPT is 2021, so it couldn’t know who won the 2023 Oscars from its internal knowledge. Therefore, it accessed external information through the browsing plugin, which allowed it to answer the question with up-to-date information. Those plugins present an integrated augmentation mechanism inside the OpenAI platform.

If you have been watching the LLM space, you might have heard of the [LangChain library](https://python.langchain.com/en/latest/index.html" \t "_blank).

**[Getting Started with LangChain: A Beginner’s Guide to Building LLM-Powered Applications](https://towardsdatascience.com/getting-started-with-langchain-a-beginners-guide-to-building-llm-powered-applications-95fc8898732c?source=post_page-----30e875d63a35--------------------------------" \t "_blank)**

[A LangChain tutorial to build anything with large language models in Python](https://towardsdatascience.com/getting-started-with-langchain-a-beginners-guide-to-building-llm-powered-applications-95fc8898732c?source=post_page-----30e875d63a35--------------------------------" \t "_blank)

[towardsdatascience.com](https://towardsdatascience.com/getting-started-with-langchain-a-beginners-guide-to-building-llm-powered-applications-95fc8898732c?source=post_page-----30e875d63a35--------------------------------" \t "_blank)

The LangChain library can be used to allow LLMs to access real-time information from various sources like Google Search, vector databases, or knowledge graphs. For example, LangChain has added a [Cypher Search chain](https://towardsdatascience.com/langchain-has-added-cypher-search-cb9d821120d5), which converts natural language questions into a Cypher statement, uses it to retrieve information from the Neo4j database, and constructs a final answer based on the provided information.

**[LangChain has added Cypher Search](https://towardsdatascience.com/langchain-has-added-cypher-search-cb9d821120d5?source=post_page-----30e875d63a35--------------------------------" \t "_blank)**

[With the LangChain library, you can conveniently generate Cypher queries, enabling an efficient retrieval of…](https://towardsdatascience.com/langchain-has-added-cypher-search-cb9d821120d5?source=post_page-----30e875d63a35--------------------------------" \t "_blank)

[towardsdatascience.com](https://towardsdatascience.com/langchain-has-added-cypher-search-cb9d821120d5?source=post_page-----30e875d63a35--------------------------------" \t "_blank)

With the Cypher Search chain, an LLM is not only used to construct a final answer but also to translate a natural language question into a Cypher query.



Cypher search in LangChain. Image by author.

Another popular library for retrieval-augmented LLM workflows is [LlamaIndex (GPT Index)](https://gpt-index.readthedocs.io/en/latest/index.html" \t "_blank). LlamaIndex is a comprehensive data framework aimed at enhancing the performance of Large Language Models (LLMs) by enabling them to leverage private or custom data.

**[LlamaIndex on TWIML AI: A Distilled Summary (using LlamaIndex)](https://medium.com/llamaindex-blog/llamaindex-on-twiml-ai-a-distilled-summary-using-llamaindex-de2a88551595?source=post_page-----30e875d63a35--------------------------------" \t "_blank)**

[Overview](https://medium.com/llamaindex-blog/llamaindex-on-twiml-ai-a-distilled-summary-using-llamaindex-de2a88551595?source=post_page-----30e875d63a35--------------------------------" \t "_blank)

[medium.com](https://medium.com/llamaindex-blog/llamaindex-on-twiml-ai-a-distilled-summary-using-llamaindex-de2a88551595?source=post_page-----30e875d63a35--------------------------------" \t "_blank)

Firstly, LlamaIndex offers data connectors that facilitate the ingestion of a variety of data sources and formats, encompassing everything from APIs, PDFs, and documents to SQL or graph data.

This feature allows for an effortless integration of existing data into the LLM. Secondly, it provides efficient mechanisms to structure the ingested data using indices and graphs, ensuring the data is suitably arranged for use with LLMs. In addition, it includes an advanced retrieval and query interface, which enables users to input an LLM prompt and receive back a context-retrieved, knowledge-augmented output.

The idea behind retrieval-augmented LLM applications like ChatGPT Plugins and LangChain is to avoid relying on internal LLM knowledge only to generate answers. Instead, LLMs are used to solve tasks like [constructing database queries from natural language](https://neo4j.com/generativeai/?utm_campaign=gen-ai&utm_content=medium&utm_source=blog) and constructing answers based on externally provided information or by utilizing plugins/agents for retrieval.

The retrieval-augmented approach has some clear advantages over the fine-tuning approach:

* The answer can cite its sources of information, which allows you to validate the information and potentially change or update the underlying information based on requirements
* Hallucinations are more unlikely to occur as you don’t rely on the internal knowledge of an LLM to answer the question and only use information that is provided in the relevant documents
* Changing, updating, and maintaining the underlying information the LLM uses is easier as you transform the problem from LLM maintenance to a database maintenance, querying and context construction problem
* Answers can be personalized based on the user context, or their access permission

On the other hand, you should consider the following limitations when using the retrieval-augmented approach:

* The answers are only as good as the smart search tool
* The application needs access to your specific knowledge base, either that be a database or other data stores
* Completely disregarding the internal knowledge of the language model limits the number of questions that can be answered
* Sometimes LLMs fail to follow instructions, so there is a risk that the context might be ignored or hallucinations occur if no relevant answer data is found in the context.

**Summary**

This blog post delves into the limitations of Large Language Models (LLMs), such as

* Knowledge cutoff,
* Hallucinations, and
* The lack of user customization.

To overcome these, we explored two concepts, namely, fine-tuning and retrieval-augmented use of LLMs.

***Fine-tuning an LLM*** involves the supervised training phase, where question-answer pairs are provided to optimize the performance of the LLM. This can be used to update and expand the LLM’s internal knowledge or fine-tune it for a specific task. However, fine-tuning fails to solve the knowledge cutoff issue as it simply pushes the cutoff to a later date. It also cannot fully eliminate hallucinations. Therefore, we recommend using the fine-tuning approach for slowly changing datasets where some hallucinations are allowed. Since fine-tuning LLMs is relatively new, we are eager to learn more about fine-tuning approaches and best practices.

The second approach to overcome the limitations of LLMs is the so-called ***retrieval-augmented generation***, where the LLM serves as a natural language interface to access external information, thereby not relying only on its internal knowledge to produce answers. Advantages of the retrieval-augmented approach include source-citing, negligible hallucinations, ease of changing and updating information, and personalization.  
However, it relies heavily on the intelligent search tool to retrieve relevant information and requires access to the user’s knowledge base. Furthermore, it can only answer queries provided it has the information required to address the question.

Keep an eye out for updates from our team as we progress the development of this project, all of which will be openly documented on our GitHub repository.

**[GitHub - neo4j/NaLLM: Repository for the NaLLM project](https://github.com/neo4j/NaLLM?source=post_page-----30e875d63a35--------------------------------" \t "_blank)**

[Welcome to the NaLLM project repository, where we are exploring and demonstrating the synergies between Neo4j and Large…](https://github.com/neo4j/NaLLM?source=post_page-----30e875d63a35--------------------------------" \t "_blank)

[github.com](https://github.com/neo4j/NaLLM?source=post_page-----30e875d63a35--------------------------------" \t "_blank)

**Read the next blog in the series:** [Multi-Hop Question Answering](https://medium.com/neo4j/knowledge-graphs-llms-multi-hop-question-answering-322113f53f51)

*/ The NaLLM Project Core Team at Neo4j: Jon Harris, Noah Mayerhofer,*

[*Oskar Hane*](https://medium.com/u/e347f94807a8?source=post_page-----30e875d63a35--------------------------------)

,

[Tomaz Bratanic](https://medium.com/u/57f13c0ea39a?source=post_page-----30e875d63a35--------------------------------)

[Neo4j](https://medium.com/tag/neo4j?source=post_page-----30e875d63a35---------------neo4j-----------------)

[Llm](https://medium.com/tag/llm?source=post_page-----30e875d63a35---------------llm-----------------)

[ChatGPT](https://medium.com/tag/chatgpt?source=post_page-----30e875d63a35---------------chatgpt-----------------)

[Chatbots](https://medium.com/tag/chatbots?source=post_page-----30e875d63a35---------------chatbots-----------------)

[Graph](https://medium.com/tag/graph?source=post_page-----30e875d63a35---------------graph-----------------)

719

7

[[](https://medium.com/@bratanic-tomaz?source=post_page-----30e875d63a35--------------------------------)](https://medium.com/@bratanic-tomaz?source=post_page-----30e875d63a35--------------------------------)

[[](https://medium.com/neo4j?source=post_page-----30e875d63a35--------------------------------)](https://medium.com/neo4j?source=post_page-----30e875d63a35--------------------------------)

[Written by Tomaz Bratanic](https://medium.com/@bratanic-tomaz?source=post_page-----30e875d63a35--------------------------------)

[7.3K Followers](https://medium.com/@bratanic-tomaz/followers?source=post_page-----30e875d63a35--------------------------------)

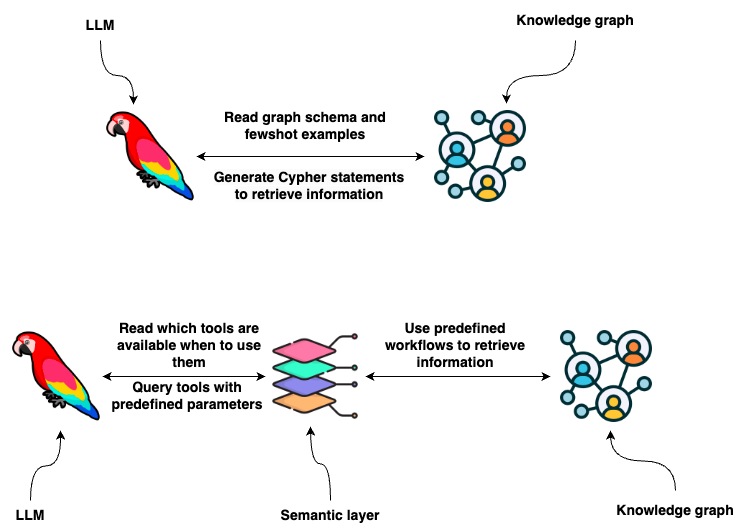
·Writer for

[Neo4j Developer Blog](https://medium.com/neo4j?source=post_page-----30e875d63a35--------------------------------)

Data explorer. Turn everything into a graph. Author of Graph algorithms for Data Science at Manning publication. <http://mng.bz/GGVN>

Follow

More from Tomaz Bratanic and Neo4j Developer Blog

[[](https://medium.com/towards-data-science/enhancing-interaction-between-language-models-and-graph-databases-via-a-semantic-layer-0a78ad3eba49?source=author_recirc-----30e875d63a35----0---------------------c7f5dd2b_7c28_450f_996c_75820e049bfb-------)](https://medium.com/towards-data-science/enhancing-interaction-between-language-models-and-graph-databases-via-a-semantic-layer-0a78ad3eba49?source=author_recirc-----30e875d63a35----0---------------------c7f5dd2b_7c28_450f_996c_75820e049bfb-------)

[[Tomaz Bratanic](https://medium.com/@bratanic-tomaz?source=author_recirc-----30e875d63a35----0---------------------c7f5dd2b_7c28_450f_996c_75820e049bfb-------)](https://medium.com/@bratanic-tomaz?source=author_recirc-----30e875d63a35----0---------------------c7f5dd2b_7c28_450f_996c_75820e049bfb-------)

[Tomaz Bratanic](https://medium.com/@bratanic-tomaz?source=author_recirc-----30e875d63a35----0---------------------c7f5dd2b_7c28_450f_996c_75820e049bfb-------)

in

[Towards Data Science](https://medium.com/towards-data-science?source=author_recirc-----30e875d63a35----0---------------------c7f5dd2b_7c28_450f_996c_75820e049bfb-------)

**[Enhancing Interaction between Language Models and Graph Databases via a Semantic Layer](https://medium.com/towards-data-science/enhancing-interaction-between-language-models-and-graph-databases-via-a-semantic-layer-0a78ad3eba49?source=author_recirc-----30e875d63a35----0---------------------c7f5dd2b_7c28_450f_996c_75820e049bfb-------)**

[Provide an LLM agent with a suite of robust tools it can use to interact with a graph database](https://medium.com/towards-data-science/enhancing-interaction-between-language-models-and-graph-databases-via-a-semantic-layer-0a78ad3eba49?source=author_recirc-----30e875d63a35----0---------------------c7f5dd2b_7c28_450f_996c_75820e049bfb-------)

[11 min read·Jan 19](https://medium.com/towards-data-science/enhancing-interaction-between-language-models-and-graph-databases-via-a-semantic-layer-0a78ad3eba49?source=author_recirc-----30e875d63a35----0---------------------c7f5dd2b_7c28_450f_996c_75820e049bfb-------)

470

[5](https://medium.com/towards-data-science/enhancing-interaction-between-language-models-and-graph-databases-via-a-semantic-layer-0a78ad3eba49?responsesOpen=true&sortBy=REVERSE_CHRON&source=author_recirc-----30e875d63a35----0---------------------c7f5dd2b_7c28_450f_996c_75820e049bfb-------)

[[](https://medium.com/neo4j/building-a-graph-llm-powered-rag-application-from-pdf-documents-24225a5baf01?source=author_recirc-----30e875d63a35----1---------------------c7f5dd2b_7c28_450f_996c_75820e049bfb-------)](https://medium.com/neo4j/building-a-graph-llm-powered-rag-application-from-pdf-documents-24225a5baf01?source=author_recirc-----30e875d63a35----1---------------------c7f5dd2b_7c28_450f_996c_75820e049bfb-------)

[[Fanghua (Joshua) Yu](https://medium.com/@yu-joshua?source=author_recirc-----30e875d63a35----1---------------------c7f5dd2b_7c28_450f_996c_75820e049bfb-------)](https://medium.com/@yu-joshua?source=author_recirc-----30e875d63a35----1---------------------c7f5dd2b_7c28_450f_996c_75820e049bfb-------)

[Fanghua (Joshua) Yu](https://medium.com/@yu-joshua?source=author_recirc-----30e875d63a35----1---------------------c7f5dd2b_7c28_450f_996c_75820e049bfb-------)

in

[Neo4j Developer Blog](https://medium.com/neo4j?source=author_recirc-----30e875d63a35----1---------------------c7f5dd2b_7c28_450f_996c_75820e049bfb-------)

**[Building A Graph+LLM Powered RAG Application from PDF Documents](https://medium.com/neo4j/building-a-graph-llm-powered-rag-application-from-pdf-documents-24225a5baf01?source=author_recirc-----30e875d63a35----1---------------------c7f5dd2b_7c28_450f_996c_75820e049bfb-------)**

[A Step-by-step Walkthrough with GenAI-Stack and OpenAI](https://medium.com/neo4j/building-a-graph-llm-powered-rag-application-from-pdf-documents-24225a5baf01?source=author_recirc-----30e875d63a35----1---------------------c7f5dd2b_7c28_450f_996c_75820e049bfb-------)

[6 min read·Jan 15](https://medium.com/neo4j/building-a-graph-llm-powered-rag-application-from-pdf-documents-24225a5baf01?source=author_recirc-----30e875d63a35----1---------------------c7f5dd2b_7c28_450f_996c_75820e049bfb-------)

344

[2](https://medium.com/neo4j/building-a-graph-llm-powered-rag-application-from-pdf-documents-24225a5baf01?responsesOpen=true&sortBy=REVERSE_CHRON&source=author_recirc-----30e875d63a35----1---------------------c7f5dd2b_7c28_450f_996c_75820e049bfb-------)

[[](https://medium.com/neo4j/why-vector-search-didnt-work-for-your-rag-solution-d95119a8c42e?source=author_recirc-----30e875d63a35----2---------------------c7f5dd2b_7c28_450f_996c_75820e049bfb-------)](https://medium.com/neo4j/why-vector-search-didnt-work-for-your-rag-solution-d95119a8c42e?source=author_recirc-----30e875d63a35----2---------------------c7f5dd2b_7c28_450f_996c_75820e049bfb-------)

[[Fanghua (Joshua) Yu](https://medium.com/@yu-joshua?source=author_recirc-----30e875d63a35----2---------------------c7f5dd2b_7c28_450f_996c_75820e049bfb-------)](https://medium.com/@yu-joshua?source=author_recirc-----30e875d63a35----2---------------------c7f5dd2b_7c28_450f_996c_75820e049bfb-------)

[Fanghua (Joshua) Yu](https://medium.com/@yu-joshua?source=author_recirc-----30e875d63a35----2---------------------c7f5dd2b_7c28_450f_996c_75820e049bfb-------)

in

[Neo4j Developer Blog](https://medium.com/neo4j?source=author_recirc-----30e875d63a35----2---------------------c7f5dd2b_7c28_450f_996c_75820e049bfb-------)

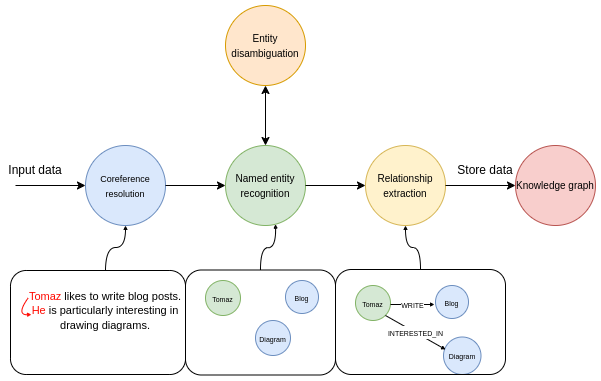
**[Why Vector Search Didn’t Work for Your RAG Solution?](https://medium.com/neo4j/why-vector-search-didnt-work-for-your-rag-solution-d95119a8c42e?source=author_recirc-----30e875d63a35----2---------------------c7f5dd2b_7c28_450f_996c_75820e049bfb-------)**

[Exploring The Shortcomings of Text Embedding Retrieval for LLM Generation](https://medium.com/neo4j/why-vector-search-didnt-work-for-your-rag-solution-d95119a8c42e?source=author_recirc-----30e875d63a35----2---------------------c7f5dd2b_7c28_450f_996c_75820e049bfb-------)

[11 min read·Dec 19, 2023](https://medium.com/neo4j/why-vector-search-didnt-work-for-your-rag-solution-d95119a8c42e?source=author_recirc-----30e875d63a35----2---------------------c7f5dd2b_7c28_450f_996c_75820e049bfb-------)

83

[1](https://medium.com/neo4j/why-vector-search-didnt-work-for-your-rag-solution-d95119a8c42e?responsesOpen=true&sortBy=REVERSE_CHRON&source=author_recirc-----30e875d63a35----2---------------------c7f5dd2b_7c28_450f_996c_75820e049bfb-------)

[[](https://medium.com/@bratanic-tomaz/constructing-knowledge-graphs-from-text-using-openai-functions-096a6d010c17?source=author_recirc-----30e875d63a35----3---------------------c7f5dd2b_7c28_450f_996c_75820e049bfb-------)](https://medium.com/@bratanic-tomaz/constructing-knowledge-graphs-from-text-using-openai-functions-096a6d010c17?source=author_recirc-----30e875d63a35----3---------------------c7f5dd2b_7c28_450f_996c_75820e049bfb-------)

[[Tomaz Bratanic](https://medium.com/@bratanic-tomaz?source=author_recirc-----30e875d63a35----3---------------------c7f5dd2b_7c28_450f_996c_75820e049bfb-------)](https://medium.com/@bratanic-tomaz?source=author_recirc-----30e875d63a35----3---------------------c7f5dd2b_7c28_450f_996c_75820e049bfb-------)

[Tomaz Bratanic](https://medium.com/@bratanic-tomaz?source=author_recirc-----30e875d63a35----3---------------------c7f5dd2b_7c28_450f_996c_75820e049bfb-------)

**[Constructing knowledge graphs from text using OpenAI functions](https://medium.com/@bratanic-tomaz/constructing-knowledge-graphs-from-text-using-openai-functions-096a6d010c17?source=author_recirc-----30e875d63a35----3---------------------c7f5dd2b_7c28_450f_996c_75820e049bfb-------)**

[Seamlessy implement information extraction pipeline with LangChain and Neo4j](https://medium.com/@bratanic-tomaz/constructing-knowledge-graphs-from-text-using-openai-functions-096a6d010c17?source=author_recirc-----30e875d63a35----3---------------------c7f5dd2b_7c28_450f_996c_75820e049bfb-------)

[11 min read·Oct 20, 2023](https://medium.com/@bratanic-tomaz/constructing-knowledge-graphs-from-text-using-openai-functions-096a6d010c17?source=author_recirc-----30e875d63a35----3---------------------c7f5dd2b_7c28_450f_996c_75820e049bfb-------)

908

[10](https://medium.com/@bratanic-tomaz/constructing-knowledge-graphs-from-text-using-openai-functions-096a6d010c17?responsesOpen=true&sortBy=REVERSE_CHRON&source=author_recirc-----30e875d63a35----3---------------------c7f5dd2b_7c28_450f_996c_75820e049bfb-------)

[See all from Tomaz Bratanic](https://medium.com/@bratanic-tomaz?source=post_page-----30e875d63a35--------------------------------)

[See all from Neo4j Developer Blog](https://medium.com/neo4j?source=post_page-----30e875d63a35--------------------------------)

Recommended from Medium

[[](https://medium.com/neo4j/construct-knowledge-graphs-from-unstructured-text-877be33300a2?source=read_next_recirc-----30e875d63a35----0---------------------43a9b80e_1f93_4a03_a918_d60631a7592d-------)](https://medium.com/neo4j/construct-knowledge-graphs-from-unstructured-text-877be33300a2?source=read_next_recirc-----30e875d63a35----0---------------------43a9b80e_1f93_4a03_a918_d60631a7592d-------)

[[Noah Mayerhofer](https://medium.com/@noah.mayerhofer?source=read_next_recirc-----30e875d63a35----0---------------------43a9b80e_1f93_4a03_a918_d60631a7592d-------)](https://medium.com/@noah.mayerhofer?source=read_next_recirc-----30e875d63a35----0---------------------43a9b80e_1f93_4a03_a918_d60631a7592d-------)

[Noah Mayerhofer](https://medium.com/@noah.mayerhofer?source=read_next_recirc-----30e875d63a35----0---------------------43a9b80e_1f93_4a03_a918_d60631a7592d-------)

in

[Neo4j Developer Blog](https://medium.com/neo4j?source=read_next_recirc-----30e875d63a35----0---------------------43a9b80e_1f93_4a03_a918_d60631a7592d-------)

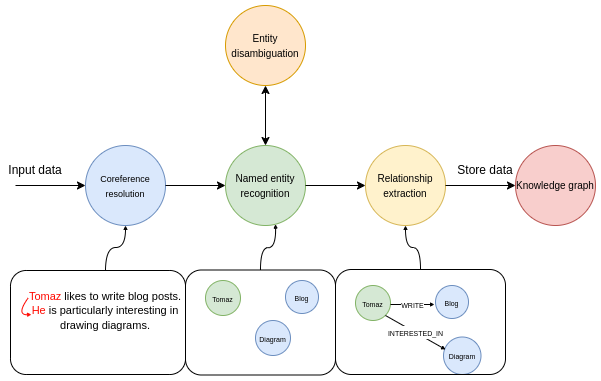
**[Construct Knowledge Graphs From Unstructured Text](https://medium.com/neo4j/construct-knowledge-graphs-from-unstructured-text-877be33300a2?source=read_next_recirc-----30e875d63a35----0---------------------43a9b80e_1f93_4a03_a918_d60631a7592d-------)**

[Our focus in this post is to extract information from unstructured text using LLMs](https://medium.com/neo4j/construct-knowledge-graphs-from-unstructured-text-877be33300a2?source=read_next_recirc-----30e875d63a35----0---------------------43a9b80e_1f93_4a03_a918_d60631a7592d-------)

[8 min read·Sep 21, 2023](https://medium.com/neo4j/construct-knowledge-graphs-from-unstructured-text-877be33300a2?source=read_next_recirc-----30e875d63a35----0---------------------43a9b80e_1f93_4a03_a918_d60631a7592d-------)

165

[4](https://medium.com/neo4j/construct-knowledge-graphs-from-unstructured-text-877be33300a2?responsesOpen=true&sortBy=REVERSE_CHRON&source=read_next_recirc-----30e875d63a35----0---------------------43a9b80e_1f93_4a03_a918_d60631a7592d-------)

[[](https://medium.com/@bratanic-tomaz/constructing-knowledge-graphs-from-text-using-openai-functions-096a6d010c17?source=read_next_recirc-----30e875d63a35----1---------------------43a9b80e_1f93_4a03_a918_d60631a7592d-------)](https://medium.com/@bratanic-tomaz/constructing-knowledge-graphs-from-text-using-openai-functions-096a6d010c17?source=read_next_recirc-----30e875d63a35----1---------------------43a9b80e_1f93_4a03_a918_d60631a7592d-------)

[[Tomaz Bratanic](https://medium.com/@bratanic-tomaz?source=read_next_recirc-----30e875d63a35----1---------------------43a9b80e_1f93_4a03_a918_d60631a7592d-------)](https://medium.com/@bratanic-tomaz?source=read_next_recirc-----30e875d63a35----1---------------------43a9b80e_1f93_4a03_a918_d60631a7592d-------)

[Tomaz Bratanic](https://medium.com/@bratanic-tomaz?source=read_next_recirc-----30e875d63a35----1---------------------43a9b80e_1f93_4a03_a918_d60631a7592d-------)

**[Constructing knowledge graphs from text using OpenAI functions](https://medium.com/@bratanic-tomaz/constructing-knowledge-graphs-from-text-using-openai-functions-096a6d010c17?source=read_next_recirc-----30e875d63a35----1---------------------43a9b80e_1f93_4a03_a918_d60631a7592d-------)**

[Seamlessy implement information extraction pipeline with LangChain and Neo4j](https://medium.com/@bratanic-tomaz/constructing-knowledge-graphs-from-text-using-openai-functions-096a6d010c17?source=read_next_recirc-----30e875d63a35----1---------------------43a9b80e_1f93_4a03_a918_d60631a7592d-------)

[11 min read·Oct 20, 2023](https://medium.com/@bratanic-tomaz/constructing-knowledge-graphs-from-text-using-openai-functions-096a6d010c17?source=read_next_recirc-----30e875d63a35----1---------------------43a9b80e_1f93_4a03_a918_d60631a7592d-------)

908

[10](https://medium.com/@bratanic-tomaz/constructing-knowledge-graphs-from-text-using-openai-functions-096a6d010c17?responsesOpen=true&sortBy=REVERSE_CHRON&source=read_next_recirc-----30e875d63a35----1---------------------43a9b80e_1f93_4a03_a918_d60631a7592d-------)

Lists

[[](https://medium.com/@m.wasalski/list/chatgpt-3742c7a4727d?source=read_next_recirc-----30e875d63a35--------------------------------)](https://medium.com/@m.wasalski/list/chatgpt-3742c7a4727d?source=read_next_recirc-----30e875d63a35--------------------------------)

[[AI-generated image of a cute tiny robot in the backdrop of ChatGPT’s logo](https://medium.com/@m.wasalski/list/chatgpt-3742c7a4727d?source=read_next_recirc-----30e875d63a35--------------------------------)](https://medium.com/@m.wasalski/list/chatgpt-3742c7a4727d?source=read_next_recirc-----30e875d63a35--------------------------------)

[[](https://medium.com/@m.wasalski/list/chatgpt-3742c7a4727d?source=read_next_recirc-----30e875d63a35--------------------------------)](https://medium.com/@m.wasalski/list/chatgpt-3742c7a4727d?source=read_next_recirc-----30e875d63a35--------------------------------)

[ChatGPT](https://medium.com/@m.wasalski/list/chatgpt-3742c7a4727d?source=read_next_recirc-----30e875d63a35--------------------------------)

[21 stories·429 saves](https://medium.com/@m.wasalski/list/chatgpt-3742c7a4727d?source=read_next_recirc-----30e875d63a35--------------------------------)

[[](https://medium.com/@MediumForTeams/list/what-is-chatgpt-7a5756752f49?source=read_next_recirc-----30e875d63a35--------------------------------)](https://medium.com/@MediumForTeams/list/what-is-chatgpt-7a5756752f49?source=read_next_recirc-----30e875d63a35--------------------------------)

[[](https://medium.com/@MediumForTeams/list/what-is-chatgpt-7a5756752f49?source=read_next_recirc-----30e875d63a35--------------------------------)](https://medium.com/@MediumForTeams/list/what-is-chatgpt-7a5756752f49?source=read_next_recirc-----30e875d63a35--------------------------------)

[[](https://medium.com/@MediumForTeams/list/what-is-chatgpt-7a5756752f49?source=read_next_recirc-----30e875d63a35--------------------------------)](https://medium.com/@MediumForTeams/list/what-is-chatgpt-7a5756752f49?source=read_next_recirc-----30e875d63a35--------------------------------)

[What is ChatGPT?](https://medium.com/@MediumForTeams/list/what-is-chatgpt-7a5756752f49?source=read_next_recirc-----30e875d63a35--------------------------------)

[9 stories·283 saves](https://medium.com/@MediumForTeams/list/what-is-chatgpt-7a5756752f49?source=read_next_recirc-----30e875d63a35--------------------------------)

[[Image by vectorjuice on FreePik](https://medium.com/@MediumStaff/list/the-new-chatbots-chatgpt-bard-and-beyond-5969c7449b7f?source=read_next_recirc-----30e875d63a35--------------------------------)](https://medium.com/@MediumStaff/list/the-new-chatbots-chatgpt-bard-and-beyond-5969c7449b7f?source=read_next_recirc-----30e875d63a35--------------------------------)

[[](https://medium.com/@MediumStaff/list/the-new-chatbots-chatgpt-bard-and-beyond-5969c7449b7f?source=read_next_recirc-----30e875d63a35--------------------------------)](https://medium.com/@MediumStaff/list/the-new-chatbots-chatgpt-bard-and-beyond-5969c7449b7f?source=read_next_recirc-----30e875d63a35--------------------------------)

[[](https://medium.com/@MediumStaff/list/the-new-chatbots-chatgpt-bard-and-beyond-5969c7449b7f?source=read_next_recirc-----30e875d63a35--------------------------------)](https://medium.com/@MediumStaff/list/the-new-chatbots-chatgpt-bard-and-beyond-5969c7449b7f?source=read_next_recirc-----30e875d63a35--------------------------------)

[The New Chatbots: ChatGPT, Bard, and Beyond](https://medium.com/@MediumStaff/list/the-new-chatbots-chatgpt-bard-and-beyond-5969c7449b7f?source=read_next_recirc-----30e875d63a35--------------------------------)

[12 stories·286 saves](https://medium.com/@MediumStaff/list/the-new-chatbots-chatgpt-bard-and-beyond-5969c7449b7f?source=read_next_recirc-----30e875d63a35--------------------------------)

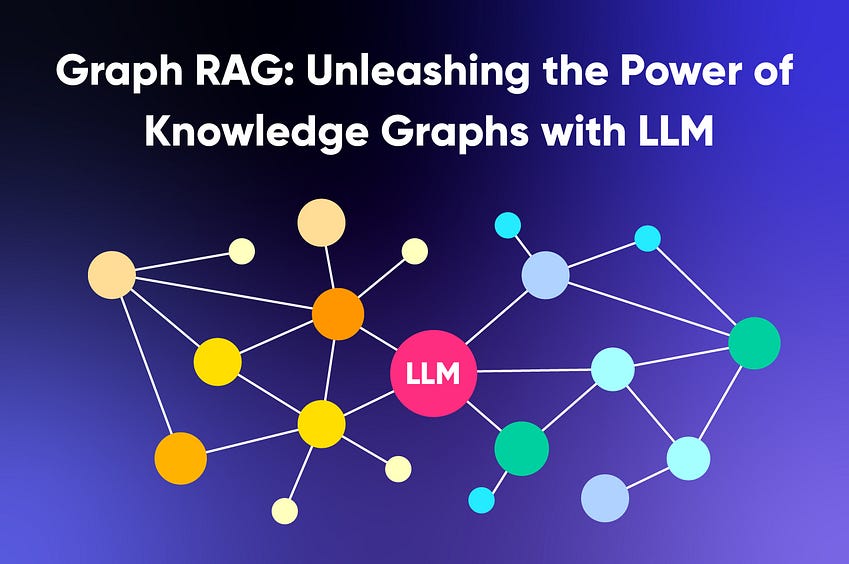
[[](https://medium.com/@nicholas.michael.janulewicz/list/chatgpt-prompts-b4c47b8e12ee?source=read_next_recirc-----30e875d63a35--------------------------------)](https://medium.com/@nicholas.michael.janulewicz/list/chatgpt-prompts-b4c47b8e12ee?source=read_next_recirc-----30e875d63a35--------------------------------)

[[](https://medium.com/@nicholas.michael.janulewicz/list/chatgpt-prompts-b4c47b8e12ee?source=read_next_recirc-----30e875d63a35--------------------------------)](https://medium.com/@nicholas.michael.janulewicz/list/chatgpt-prompts-b4c47b8e12ee?source=read_next_recirc-----30e875d63a35--------------------------------)

[[](https://medium.com/@nicholas.michael.janulewicz/list/chatgpt-prompts-b4c47b8e12ee?source=read_next_recirc-----30e875d63a35--------------------------------)](https://medium.com/@nicholas.michael.janulewicz/list/chatgpt-prompts-b4c47b8e12ee?source=read_next_recirc-----30e875d63a35--------------------------------)

[ChatGPT prompts](https://medium.com/@nicholas.michael.janulewicz/list/chatgpt-prompts-b4c47b8e12ee?source=read_next_recirc-----30e875d63a35--------------------------------)

[36 stories·1054 saves](https://medium.com/@nicholas.michael.janulewicz/list/chatgpt-prompts-b4c47b8e12ee?source=read_next_recirc-----30e875d63a35--------------------------------)

[[](https://medium.com/@nebulagraph/graph-rag-the-new-llm-stack-with-knowledge-graphs-e1e902c504ed?source=read_next_recirc-----30e875d63a35----0---------------------43a9b80e_1f93_4a03_a918_d60631a7592d-------)](https://medium.com/@nebulagraph/graph-rag-the-new-llm-stack-with-knowledge-graphs-e1e902c504ed?source=read_next_recirc-----30e875d63a35----0---------------------43a9b80e_1f93_4a03_a918_d60631a7592d-------)

[[NebulaGraph Database](https://medium.com/@nebulagraph?source=read_next_recirc-----30e875d63a35----0---------------------43a9b80e_1f93_4a03_a918_d60631a7592d-------)](https://medium.com/@nebulagraph?source=read_next_recirc-----30e875d63a35----0---------------------43a9b80e_1f93_4a03_a918_d60631a7592d-------)

[NebulaGraph Database](https://medium.com/@nebulagraph?source=read_next_recirc-----30e875d63a35----0---------------------43a9b80e_1f93_4a03_a918_d60631a7592d-------)

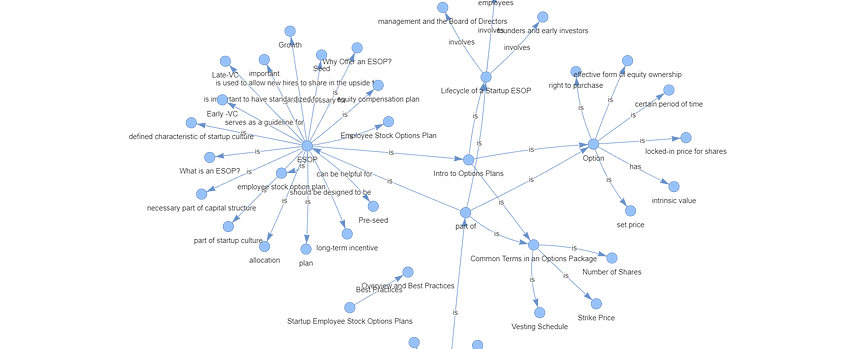
**[Graph RAG: Unleashing the Power of Knowledge Graphs with LLM](https://medium.com/@nebulagraph/graph-rag-the-new-llm-stack-with-knowledge-graphs-e1e902c504ed?source=read_next_recirc-----30e875d63a35----0---------------------43a9b80e_1f93_4a03_a918_d60631a7592d-------)**

[In the era of information overload, sifting through vast amounts of data to provide accurate search results in an engaging and…](https://medium.com/@nebulagraph/graph-rag-the-new-llm-stack-with-knowledge-graphs-e1e902c504ed?source=read_next_recirc-----30e875d63a35----0---------------------43a9b80e_1f93_4a03_a918_d60631a7592d-------)

[6 min read·Sep 8, 2023](https://medium.com/@nebulagraph/graph-rag-the-new-llm-stack-with-knowledge-graphs-e1e902c504ed?source=read_next_recirc-----30e875d63a35----0---------------------43a9b80e_1f93_4a03_a918_d60631a7592d-------)

240

[1](https://medium.com/@nebulagraph/graph-rag-the-new-llm-stack-with-knowledge-graphs-e1e902c504ed?responsesOpen=true&sortBy=REVERSE_CHRON&source=read_next_recirc-----30e875d63a35----0---------------------43a9b80e_1f93_4a03_a918_d60631a7592d-------)

[[](https://medium.com/dphi-tech/implement-rag-with-knowledge-graph-and-llama-index-6a3370e93cdd?source=read_next_recirc-----30e875d63a35----1---------------------43a9b80e_1f93_4a03_a918_d60631a7592d-------)](https://medium.com/dphi-tech/implement-rag-with-knowledge-graph-and-llama-index-6a3370e93cdd?source=read_next_recirc-----30e875d63a35----1---------------------43a9b80e_1f93_4a03_a918_d60631a7592d-------)

[[Plaban Nayak](https://medium.com/@nayakpplaban?source=read_next_recirc-----30e875d63a35----1---------------------43a9b80e_1f93_4a03_a918_d60631a7592d-------)](https://medium.com/@nayakpplaban?source=read_next_recirc-----30e875d63a35----1---------------------43a9b80e_1f93_4a03_a918_d60631a7592d-------)

[Plaban Nayak](https://medium.com/@nayakpplaban?source=read_next_recirc-----30e875d63a35----1---------------------43a9b80e_1f93_4a03_a918_d60631a7592d-------)

in

[AI Planet](https://medium.com/dphi-tech?source=read_next_recirc-----30e875d63a35----1---------------------43a9b80e_1f93_4a03_a918_d60631a7592d-------)

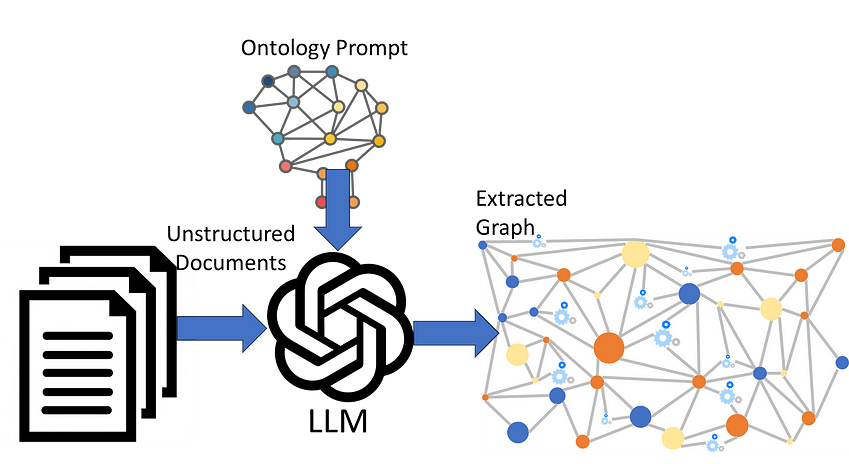
**[Implement RAG with Knowledge Graph and Llama-Index](https://medium.com/dphi-tech/implement-rag-with-knowledge-graph-and-llama-index-6a3370e93cdd?source=read_next_recirc-----30e875d63a35----1---------------------43a9b80e_1f93_4a03_a918_d60631a7592d-------)**

[Hallucination is a common problem when working with large language models (LLMs). LLMs generate fluent and coherent text but often generate…](https://medium.com/dphi-tech/implement-rag-with-knowledge-graph-and-llama-index-6a3370e93cdd?source=read_next_recirc-----30e875d63a35----1---------------------43a9b80e_1f93_4a03_a918_d60631a7592d-------)

[25 min read·Dec 3, 2023](https://medium.com/dphi-tech/implement-rag-with-knowledge-graph-and-llama-index-6a3370e93cdd?source=read_next_recirc-----30e875d63a35----1---------------------43a9b80e_1f93_4a03_a918_d60631a7592d-------)

692

[7](https://medium.com/dphi-tech/implement-rag-with-knowledge-graph-and-llama-index-6a3370e93cdd?responsesOpen=true&sortBy=REVERSE_CHRON&source=read_next_recirc-----30e875d63a35----1---------------------43a9b80e_1f93_4a03_a918_d60631a7592d-------)

[[](https://medium.com/gopenai/llm-ontology-prompting-for-knowledge-graph-extraction-efdcdd0db3a1?source=read_next_recirc-----30e875d63a35----2---------------------43a9b80e_1f93_4a03_a918_d60631a7592d-------)](https://medium.com/gopenai/llm-ontology-prompting-for-knowledge-graph-extraction-efdcdd0db3a1?source=read_next_recirc-----30e875d63a35----2---------------------43a9b80e_1f93_4a03_a918_d60631a7592d-------)

[[Peter Lawrence, answering users' data questions](https://medium.com/@peter.lawrence_47665?source=read_next_recirc-----30e875d63a35----2---------------------43a9b80e_1f93_4a03_a918_d60631a7592d-------)](https://medium.com/@peter.lawrence_47665?source=read_next_recirc-----30e875d63a35----2---------------------43a9b80e_1f93_4a03_a918_d60631a7592d-------)

[Peter Lawrence, answering users' data questions](https://medium.com/@peter.lawrence_47665?source=read_next_recirc-----30e875d63a35----2---------------------43a9b80e_1f93_4a03_a918_d60631a7592d-------)

in

[GoPenAI](https://medium.com/gopenai?source=read_next_recirc-----30e875d63a35----2---------------------43a9b80e_1f93_4a03_a918_d60631a7592d-------)

**[LLM Ontology-prompting for Knowledge Graph Extraction](https://medium.com/gopenai/llm-ontology-prompting-for-knowledge-graph-extraction-efdcdd0db3a1?source=read_next_recirc-----30e875d63a35----2---------------------43a9b80e_1f93_4a03_a918_d60631a7592d-------)**

[Prompting an LLM with an ontology to drive Knowledge Graph extraction from unstructured documents](https://medium.com/gopenai/llm-ontology-prompting-for-knowledge-graph-extraction-efdcdd0db3a1?source=read_next_recirc-----30e875d63a35----2---------------------43a9b80e_1f93_4a03_a918_d60631a7592d-------)

[5 min read·Jun 27, 2023](https://medium.com/gopenai/llm-ontology-prompting-for-knowledge-graph-extraction-efdcdd0db3a1?source=read_next_recirc-----30e875d63a35----2---------------------43a9b80e_1f93_4a03_a918_d60631a7592d-------)

621

[9](https://medium.com/gopenai/llm-ontology-prompting-for-knowledge-graph-extraction-efdcdd0db3a1?responsesOpen=true&sortBy=REVERSE_CHRON&source=read_next_recirc-----30e875d63a35----2---------------------43a9b80e_1f93_4a03_a918_d60631a7592d-------)

[[](https://medium.com/@yu-joshua/why-vectors-should-be-stored-together-with-knowledge-graph-41e4280e2f02?source=read_next_recirc-----30e875d63a35----3---------------------43a9b80e_1f93_4a03_a918_d60631a7592d-------)](https://medium.com/@yu-joshua/why-vectors-should-be-stored-together-with-knowledge-graph-41e4280e2f02?source=read_next_recirc-----30e875d63a35----3---------------------43a9b80e_1f93_4a03_a918_d60631a7592d-------)

[[Fanghua (Joshua) Yu](https://medium.com/@yu-joshua?source=read_next_recirc-----30e875d63a35----3---------------------43a9b80e_1f93_4a03_a918_d60631a7592d-------)](https://medium.com/@yu-joshua?source=read_next_recirc-----30e875d63a35----3---------------------43a9b80e_1f93_4a03_a918_d60631a7592d-------)

[Fanghua (Joshua) Yu](https://medium.com/@yu-joshua?source=read_next_recirc-----30e875d63a35----3---------------------43a9b80e_1f93_4a03_a918_d60631a7592d-------)

**[Why Vectors Should Be Stored Together with Knowledge Graph?](https://medium.com/@yu-joshua/why-vectors-should-be-stored-together-with-knowledge-graph-41e4280e2f02?source=read_next_recirc-----30e875d63a35----3---------------------43a9b80e_1f93_4a03_a918_d60631a7592d-------)**

[An architecture review of Neo4j as both vector and graph store](https://medium.com/@yu-joshua/why-vectors-should-be-stored-together-with-knowledge-graph-41e4280e2f02?source=read_next_recirc-----30e875d63a35----3---------------------43a9b80e_1f93_4a03_a918_d60631a7592d-------)

[8 min read·Oct 11, 2023](https://medium.com/@yu-joshua/why-vectors-should-be-stored-together-with-knowledge-graph-41e4280e2f02?source=read_next_recirc-----30e875d63a35----3---------------------43a9b80e_1f93_4a03_a918_d60631a7592d-------)

117

[1](https://medium.com/@yu-joshua/why-vectors-should-be-stored-together-with-knowledge-graph-41e4280e2f02?responsesOpen=true&sortBy=REVERSE_CHRON&source=read_next_recirc-----30e875d63a35----3---------------------43a9b80e_1f93_4a03_a918_d60631a7592d-------)

[See more recommendations](https://medium.com/?source=post_page-----30e875d63a35--------------------------------)

[Help](https://help.medium.com/hc/en-us?source=post_page-----30e875d63a35--------------------------------)

[Status](https://medium.statuspage.io/?source=post_page-----30e875d63a35--------------------------------)

[About](https://medium.com/about?autoplay=1&source=post_page-----30e875d63a35--------------------------------)

[Careers](https://medium.com/jobs-at-medium/work-at-medium-959d1a85284e?source=post_page-----30e875d63a35--------------------------------)

[Blog](https://blog.medium.com/?source=post_page-----30e875d63a35--------------------------------)

[Privacy](https://policy.medium.com/medium-privacy-policy-f03bf92035c9?source=post_page-----30e875d63a35--------------------------------)

[Terms](https://policy.medium.com/medium-terms-of-service-9db0094a1e0f?source=post_page-----30e875d63a35--------------------------------)

[Text to speech](https://speechify.com/medium?source=post_page-----30e875d63a35--------------------------------)

[Teams](https://medium.com/business?source=post_page-----30e875d63a35--------------------------------)