LLM Analysis of 10-K and 10-Q Filings: RAG Results

[1]Shyam Balagurumurthy Viswanathan, [2]Amritha Arun Babu Mysore, [3]Arjun Karat, [4]Raghavan Muthuregunathan

[1]Integrity Science Engineer, [2]ML Senior Product Manager, [3]Senior Member of Technical Staff
Senior Engineering Manager
City University of New York
New York, USA.

***Abstract*- In the field of financial analytics, Large Language Models (LLMs) have become significant tools for analyzing and interpreting complex, unstructured text, like the content found in financial reports. This paper focuses on the use of LLMs to study 10-K and 10-Q filings, which are essential financial statements required from publicly traded companies. These reports thoroughly view a company's financial situation and future outlook, targeting a broad audience, including investors, analysts, regulators, employees, and the general public. The main part of this research involves combining Retrieval-Augmented Generation (RAG) models and knowledge graphs with LLMs to improve the analysis of these financial documents. A key challenge in understanding 10-K and 10-Q reports is figuring out which companies are competitors and how they fit into the market, something that regular LLM methods don't fully address. Our approach uses RAG models and knowledge graphs to overcome this issue, offering a better understanding of how companies compete and where they stand in their industry based on these reports.**

This study aims to show that integrating LLMs with RAG and knowledge graphs can significantly improve how we analyze financial data, leading to a deeper and more complete understanding of a company's position in the market, its competition, and potential risks. This research goes beyond just analyzing financial reports, suggesting a new way to use AI in business intelligence and market research.

INTRODUCTION

In the complex and dynamic world of financial investment, a deep understanding of intricate financial documents is crucial for success. This is particularly true for analyzing 10-K and 10-Q filings – comprehensive reports that public companies must publish annually and quarterly. These filings have essential financial data and insights critical for making informed investment decisions. However, the challenge lies in the fact that these reports are often complex and technical, making them difficult to decipher for those who do not have a specialized financial background. This barrier can prevent many potential investors from accessing valuable insights and hinder their ability to make strategic, informed investment decisions.

Furthermore, analyzing these reports is not only about understanding the raw financial data but also about interpreting it in the context of broader market conditions. This includes connecting the dots between the data in these reports and global economic indicators, geopolitical events, and fluctuations in raw material markets. Such a comprehensive analysis requires financial expertise and a keen understanding of global market dynamics, which adds a layer of complexity and often leads to a cognitive delay in understanding the complete financial picture.

Recognizing this gap, there is a significant opportunity to democratize access to these insights, making them available and understandable to a wider and more diverse range of investors. This democratization is not just about increasing transparency; it's about leveling the playing field and giving all investors, regardless of their financial acumen, the tools to make well-informed decisions. This approach could open the investment world to new participants, fostering a more inclusive financial environment.

Our research has led us to the potential solution of an Artificial Intelligence (AI) system based on Large Language Models (LLMs). These advanced AI models can transform the way investors interact with financial data. By providing an intuitive, natural language interface, LLMs can assist users in extracting and interpreting complex financial information from 10-K and 10-Q filings. This system can answer queries, offer insights, and even provide contextual analysis by integrating broader economic trends, industry developments, and competitor innovations into its responses. Such a tool would not only make financial data more accessible but also ensure that the insights derived are comprehensive and relevant to the current market scenario. This AI-driven approach promises to enhance the investment process, making it more intuitive and insightful for investors and financial analysts alike, regardless of their background in finance.

LITERATURE REVIEW

This paper examines three key areas: 10-K and 10-Q filings, Retrieval-Augmented Generation (RAG), and integrating knowledge graphs with RAG. These components form the foundation of our research on enhancing financial document analysis using advanced AI techniques.

10-K and 10-Q Filings

The 10-K and 10-Q filings[1] are pivotal documents in the financial sector, providing comprehensive overviews of a company's financial performance. The 10-K[12] is an annual report that offers a detailed summary of a company's financial activity over the past year, while the 10-Q[13] is a quarterly report with similar, though less detailed, information. Papers in this area primarily focus on the importance of these filings for stakeholders, including investors, analysts, and regulators. Studies have highlighted how these documents contain critical information on a company's financial health, market position/competition[2][3], risk factors, market dynamics, and management's perspective on the future[4]. However, the complexity and volume of data in these reports often make them challenging to analyze, necessitating advanced analytical tools to extract meaningful insights.

Retrieval-Augmented Generation (RAG)

Retrieval-augmented generation [5]is a recent development in the field of Generative AI[6]. It combines the capabilities of generative large language models with external information retrieval, allowing the model to access a vast repository of information beyond its training data. This approach has been particularly effective in enhancing the performance of language models in tasks requiring detailed, fact-based responses. In financial analysis, RAG models can be instrumental in parsing and understanding complex documents[7], such as 10-K and 10-Q filings. The literature shows that RAG models improve the accuracy of information extraction and enable the contextualization of data by linking to relevant external information sources. It also allows us to provide necessary information externally.

Knowledge Graphs with RAG

Integrating knowledge graphs with RAG[8] models represents an innovative data processing and analysis approach. Knowledge graphs are structured representations of facts and relationships[9], providing a framework for organizing and contextualizing information. Combined with RAG models, they allow for a more nuanced understanding of complex datasets. In financial analysis, this integration can significantly enhance the ability to interpret dense financial documents into nodes and edges[10]. The existing literature underscores the potential of this combination in providing a holistic view of a company's financial status by not only extracting data from the filings but also connecting it with more comprehensive economic indicators and market trends. This holistic approach is instrumental in offering deeper insights and more accurate predictions based on financial reports.

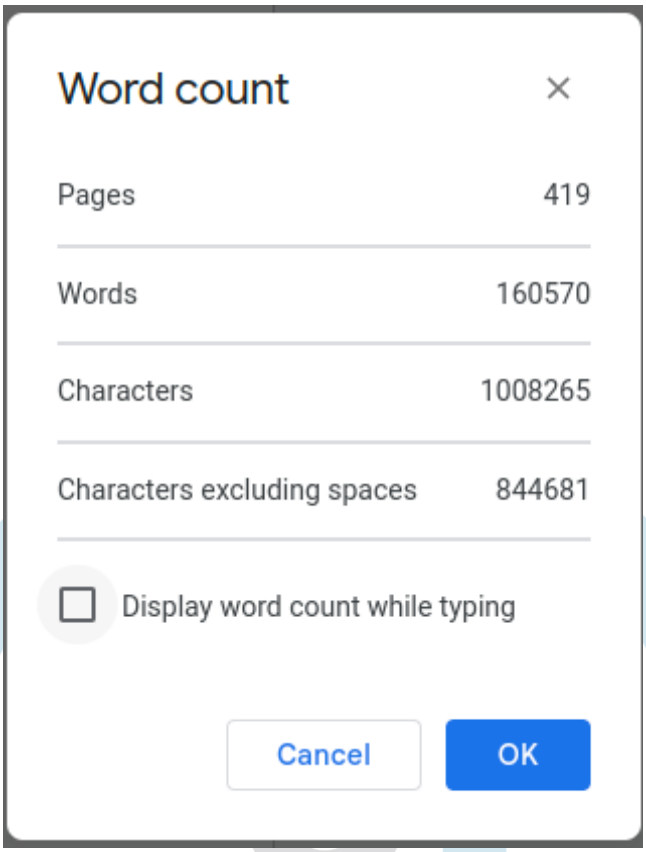
Together, these three areas form a comprehensive framework for understanding the current state of financial document analysis and the potential advancements through AI and NLP technologies. This review lays the groundwork for exploring how LLMs, augmented by RAG and knowledge graphs, can revolutionize the analysis of 10-K and 10-Q filings, making them more accessible and insightful for a broader range of users. Using RAG models and knowledge graphs, financial analysts can leverage the power of LLMs to extract and contextualize information from complex financial documents, enabling more accurate analysis and informed decision-making.

Analyzing Financial Filings

LLMs to Analyze 10-K and 10-Q Filings

1. LLMs to Analyze 10-K and 10-Q Filings

Large Language Models (LLMs) like GPT-3.5, GPT-4, LLAMA, BARD, and other LLM models are revolutionary in processing natural language. We can input the financial documents, such as 10-K and 10-Q filings, to generate a summary, ask questions on the text, perform entity extraction, and perform other analytical activities on the document. We have taken one particular company's filings to the SEC for this research. One 10-Q filing is 419 page.



Context window is limited for many of the LLM models. With the above word count, it will not work. GPT-4 has a higher context window with large document-size uploads. It answers many of the questions from the filings correctly. However, it needed to have all the information about the previous filings, the board of directors, and other specific details about the company.

You
board of directors of the company?

ChatGPT
The Board of Directors of Duck Creek Technologies is significantly influenced by Apax VIII Fund and Accenture plc. These entities have a notable impact on the composition of the board, as well as on the company's management, business plans, and policies. This influence also extends to potential conflicts of interest between Apax and Accenture and the other stockholders of Duck Creek Technologies [,,].

However, the specific names of the individual members of the Board of Directors are not provided in the document. For the most current information regarding the members of the Board of Directors, it would be best to refer to the company's official website or their most recent investor relations communications.

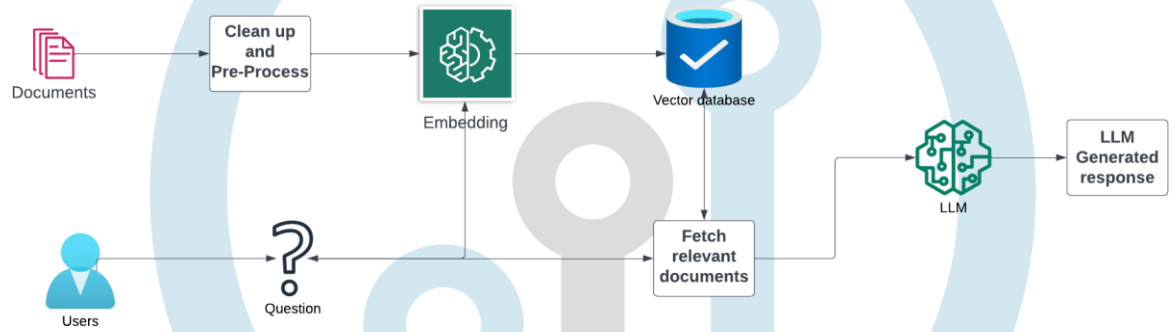
You
headquarters of the company

ChatGPT
The headquarters of Duck Creek Technologies LLC is located at 161 North Clark Street, Chicago, Illinois 60601 [,,].

Cons: In the above screenshot, information about the board of directors is unknown, and the headquarters needs to be more accurate. LLMs tend to hallucinate and provide an address that is not accurate. In short, despite their impressive capabilities, models have inherent limitations due to their context window size. The performance of these models exhibits a U-shaped curve, with optimal outcomes when relevant information is at the beginning or end of the context window. When essential information is in the middle of a lengthy text, these models need help [11]. This limitation significantly impacts their ability to process and interpret long, complex documents like 10-K and 10-Q filings. The challenge arises because LLMs, in their standard form, can only refer to a limited scope of text for generating responses. Therefore, when faced with extensive documents, they might need more critical insights or connections in the text, leading to incomplete or inaccurate analysis.

2. RAG addresses the context window with VectorDB

Retrieval-augmented generation (RAG) models address the limitations of LLMs by integrating external information retrieval capabilities. These models use vector databases to fetch relevant external data, extending the LLMs context window. RAG systems' flexibility allows various retrieval strategies, such as cosine similarity and BM25, to enhance accuracy. This approach not only broadens the information base accessible to the LLM but also improves the relevance of the generated responses. By leveraging external databases, RAG models can supplement the LLM's understanding with up-to-date and comprehensive information, making them particularly effective for tasks requiring detailed, fact-based responses.



For the current work, we will use the vector called Pinecone. We can choose any embedding model. For embedding the 10-Q financial document, we will select OpenAI embeddings. We have also fed additional information, such as the board of directors and their website information, into the RAG.

```
board of directors of the company?

 The board of directors of Duck Creek Technologies includes Bill Bloom (Independ
ent Director), Frank Pelzer (Independent Director), Jason Wright (Chairman), Jul
ie Dodd (Director), Kathy Crusco (Independent Director), Roy Mackenzie (Director
), Stuart D. Nicoll (Director), and Sunil Rajasekar (Director).

Headquarters of the company?

 The headquarters of Duck Creek Technologies LLC is located at 161 North Clark S
treet, Chicago, IL 60601.
```

It is also possible to feed future filings and all the company financial reports in the vector database and search across different documents. By accessing a vast repository of external information and knowledge, the RAG model can provide a comprehensive and detailed understanding of complex financial reports, enabling more accurate analysis and informed decision-making. These enhancements further improve the RAG model's ability to adapt to changing domains and ensure that it remains relevant and effective in processing and analyzing the latest financial reports and filings.

With the evolving landscape of financial reporting and the dynamic nature of business environments, the ability to update and finetune the RAG model with the most recent information is crucial for maintaining its accuracy and relevance. These advancements in RAG-end2end ensure that the model remains up-to-date and contribute to its capability to provide precise and contextualized insights from complex financial documents.

Limitations of RAG with VectorDB solutions: The application of Retrieval-Augmented Generation (RAG) with vector databases like Pinecone in analyzing financial documents, such as 10-K and 10-Q filings, presents a few challenges:

1. Accuracy of Data in RAG: The accuracy of an RAG model's responses heavily depends on the quality of the data in the vector database. If this data needs to be more accurate or updated, it can lead to incorrect conclusions or insights.
2. Challenges with Legacy Information: Removing outdated or irrelevant information from the vector database can take time and effort. This can result in the RAG model accessing and using legacy data that may not be applicable to current analyses.
3. Need for Comprehensive Data Loading: For effective performance, RAG models require loading extensive contextual data into the vector database. This can be resource-intensive and may limit the model's practicality in specific scenarios.
4. Missing Entity Relationships: RAG models might not effectively capture or understand the relationships between different entities. This can be a significant drawback in financial analysis, where understanding the interconnections between various financial elements is crucial.
5. Conflicting Information: Conflicting data in the vector database can complicate the analysis, leading to ambiguous or contradictory insights.

These limitations highlight the need for careful data management and the development of more sophisticated RAG models to leverage their potential in financial document analysis fully.

Different vectorDB in the market

Several vector databases are available in the market, each designed to support different aspects of data retrieval and management. Some notable vector databases include Pinecone, FAISS, Weaviate, and Chroma. These databases vary in features and capabilities, catering to different requirements in terms of scalability, efficiency, and data handling. For instance, some vector databases are optimized for large-scale data retrieval, while others focus on providing more efficient indexing and faster query responses. The choice of a suitable vector database is crucial for the optimal performance of RAG models, as it directly impacts the quality and speed of information retrieval.

3. Integration of knowledge graphs with RAGs

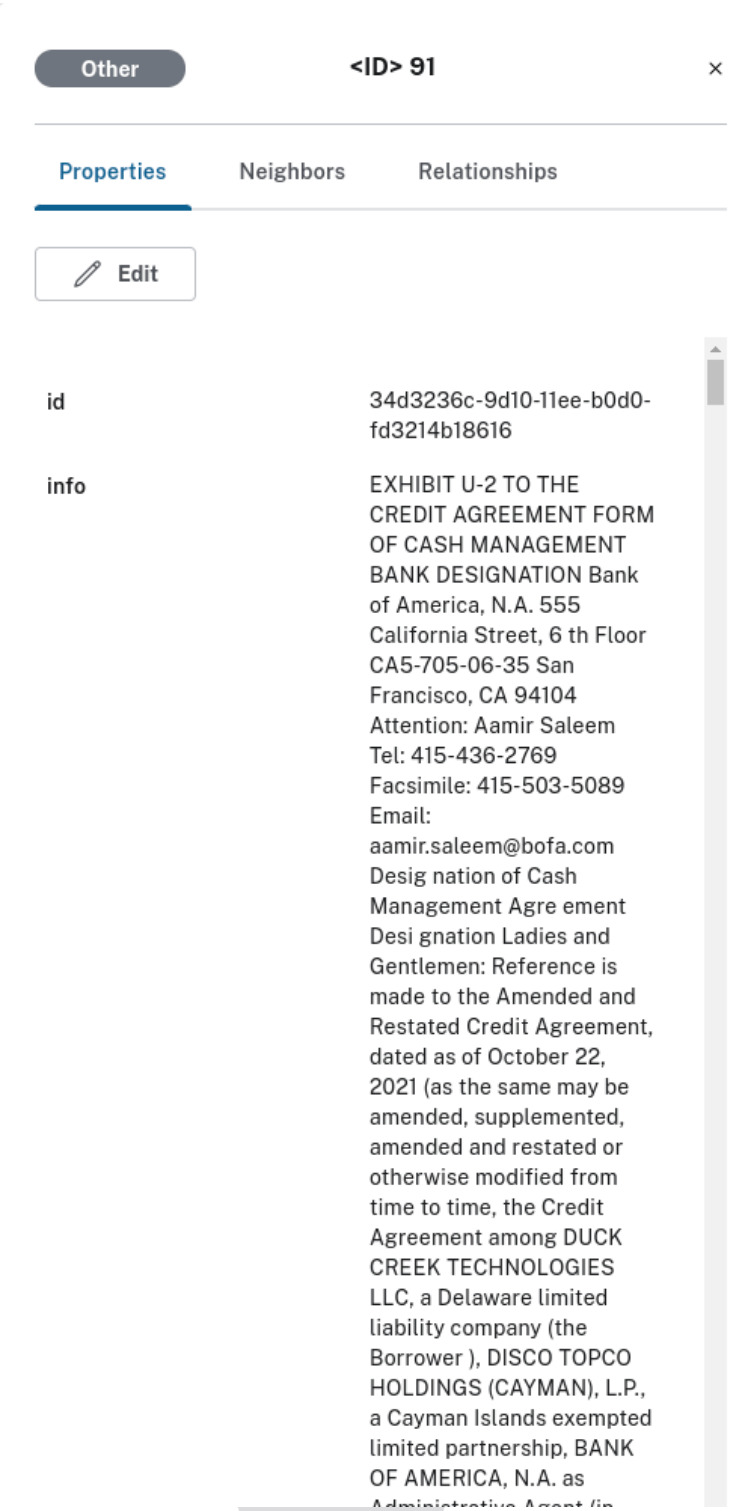
Integrating knowledge graphs with Retrieval-Augmented Generation (RAG) models represents a significant advancement in data analysis, particularly for complex documents such as financial reports. Knowledge graphs offer a structured way to express and organize information, encompassing a wide range of entities and their relationships. This structured representation is crucial in providing context and depth to the data, enabling the RAG models to understand and interpret complex relationships and dependencies more effectively.

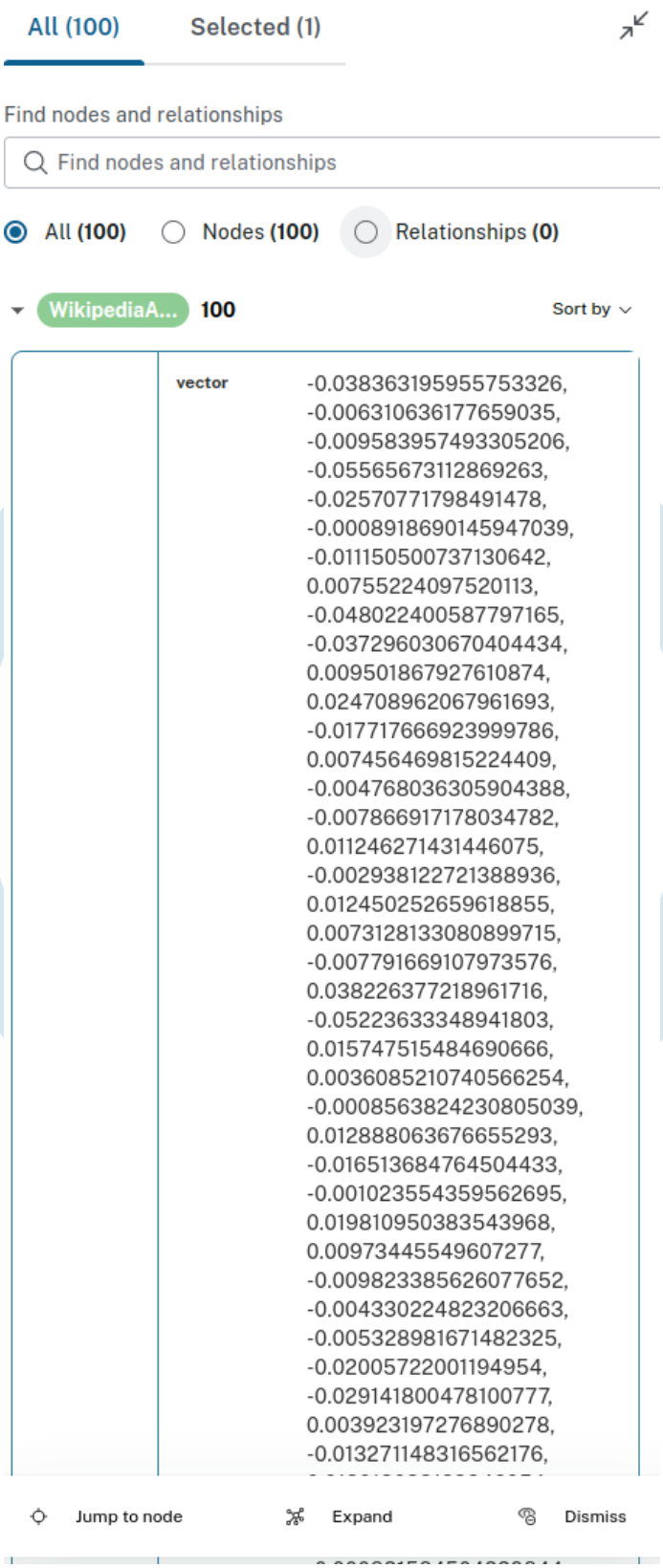
In financial document analysis, this integration is particularly valuable. Financial reports, such as 10-K and 10-Q filings, are dense with information and often contain intricate relationships between financial metrics, market trends, and company strategies. Knowledge graphs help in mapping out these relationships, making it easier for RAG models to navigate through the complexities of these documents. This results in a more nuanced and accurate analysis, as the models can better understand the context and significance of each piece of information.

Moreover, knowledge graphs facilitate a more dynamic and adaptive approach to data analysis. As new data becomes available, the knowledge graph can be updated, allowing the RAG model to refine its understanding and analysis of the subject matter continually. This is particularly important in the ever-evolving landscape of financial markets, where timely and accurate information is crucial for decision-making.

The integration also enhances the ability of RAG models to handle unstructured data. Financial documents often contain structured and unstructured data, including textual narratives, numerical data, and complex financial metrics. Knowledge graphs assist in structuring this unstructured data, thereby improving the model's ability to process and analyze it effectively.

For this research, we will use Neo4J; we have processed the entire 10-Q financial in nodes, and the embeddings of that text are added to it.


Other
<ID> 91
Properties
Neighbors
Relationships
Edit
id
34d3236c-9d10-11ee-b0d0-fd3214b18616
info
EXHIBIT U-2 TO THE CREDIT AGREEMENT FORM OF CASH MANAGEMENT BANK DESIGNATION Bank of America, N.A. 555 California Street, 6 th Floor CA5-705-06-35 San Francisco, CA 94104 Attention: Aamir Saleem Tel: 415-436-2769 Facsimile: 415-503-5089 Email: aamir.saleem@bofa.com Desig nation of Cash Management Agre ement Desi gnation Ladies and Gentlemen: Reference is made to the Amended and Restated Credit Agreement, dated as of October 22, 2021 (as the same may be amended, supplemented, amended and restated or otherwise modified from time to time, the Credit Agreement among DUCK CREEK TECHNOLOGIES LLC, a Delaware limited liability company (the Borrower), DISCO TOPCO HOLDINGS (CAYMAN), L.P., a Cayman Islands exempted limited partnership, BANK OF AMERICA, N.A. as



Using LLMs augmented by RAG and knowledge graphs offers substantial benefits in the analysis of financial filings. This approach enables a more comprehensive and accurate understanding of complex documents like 10-K and 10-Q filings. It facilitates the extraction of detailed insights and contextualization within the broader financial landscape, leading to more informed and insightful analyses. This augmented approach is particularly beneficial for stakeholders who require a deep understanding of a company's financial status, market position, and potential risks.

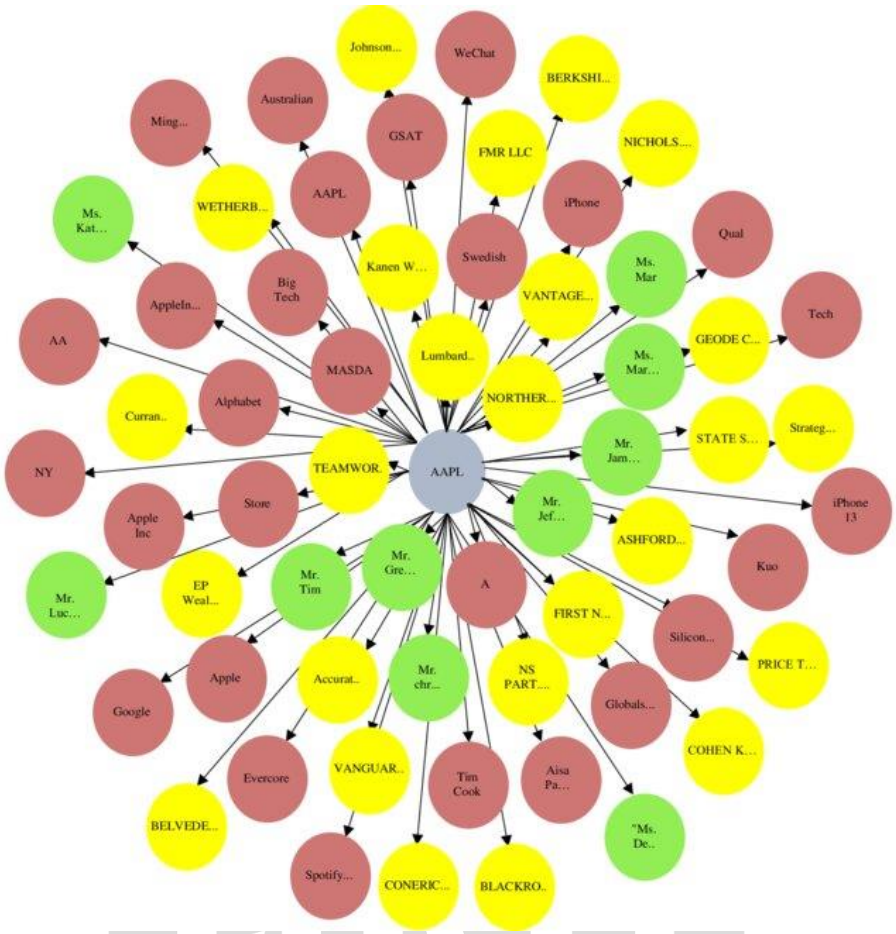
Different knowledge graph databases in the market

Various knowledge graph databases offer unique functionalities, like Neo4j, AWS Neptune, ArangoDB, TigerGraph, and Apache Giraph. These databases differ in handling graph-structured data, scalability, and ease of integration with

other systems. Selecting the appropriate database depends on the specific requirements of the task at hand, such as the dataset's size, the queries' complexity, and the desired performance metrics.

Entity Extraction from the documents and creating relationships

Entity extraction and establishing relationships between them are crucial to analyzing financial documents. Using LLMs, RAG, and knowledge graphs efficiently identifies critical entities such as companies, products, and market trends. More importantly, these technologies aid in understanding the relationships and interdependencies among these entities, offering a more straightforward and comprehensive view of a company's financial position and the broader market landscape.



Relationship Establishment

In addition to entity extraction, knowledge graphs enable robust relationships between the identified entities. Understanding the intricate dependencies and interactions between various financial elements is pivotal for comprehensive financial analysis. Knowledge graphs provide a framework for mapping out these relationships, offering a holistic view of how different entities influence and relate to one another within financial documents.

Further Advancements in Entity Extraction and Relationship Mapping

Additionally, advancements in natural language processing and knowledge graph technologies continue to enhance the accuracy and precision of entity extraction and relationship establishment.

Domains where AI Financial Filing analysis such as 10-K and 10-Q analysis can be useful

LLM analysis of corporate filings like 10-Ks and 10-Qs can provide valuable insights across many sectors, below are some such example usecase1:

1. AI Financial Filing analysis can be leveraged by small investment firms to enhance research and identify promising investment opportunities across multiple markets in minutes compared to days. The automated analysis saves time compared to manual review.
2. Economists around the world or at institutions can use AI Financial Filing analysis to summarize, understand the first level of information to get bigger picture of company performance, emerging trends, and risks across industries easily. This enables the economists to choose the areas they would like to deep dive to form a more accurate economic forecasting and policy recommendations.
3. Banking institutions can use the AI Financial filings analysis to assess lending and credit risks this process will enable them to make more informed decisions around capital allocation.

4. Individual investors now have access to professional or institutional-level insights without the need for expertise in finance or Wall Street expertise. The AI levels the playing field for retail investors looking to pick stocks aligned with their strategy. This also makes it more affordable and democratizes learning of financial analysis.
5. Public companies can benchmark against peers, identify improvement opportunities, and stay on top of emerging risks ahead of time based on the AI's broad analysis.

LLM-based AI financial filing analysis unlocks value in textual filings that can benefit everyone from individuals to economists to corporations. This data analysis augments existing decision making across the financial sector. One such example is empowering Investors through AI-Driven Research will reduce time and enable high impactful stock investment decisions.

Here is an example of - John, an individual retail investor looking to pick stocks for his retirement portfolio uses AI financial filing analysis to make informed decision.Analyzing company filings and piecing together an investment thesis is time-consuming. This is where an AI-powered solution can help level the playing field.

The AI financial filing analysis leverages large language models and retrieval augmented generation, to rapidly digest and summarize key data from 10-K and 10-Q filings. John can use this tool to specify his investing goals and criteria. The AI solution returns a shortlist of companies with growth potential along with executive summaries covering financials, business risks, and competitive strategy.

With just a few hours spent reviewing the AI-generated research instead of days poring through dense filings, John has high-level insights to make informed investing decisions. The AI assistant saves time while personalizing discoveries to John's strategy.

This use case highlights how AI can empower retail investors. Individuals now have an automated way to turn complex regulatory filings into accessible and actionable analysis. AI levels the asymmetric information advantages institutions wield over individual investors. By democratizing access to institutional-caliber insights, it improves capital allocation and financial outcomes for millions of retail investors like John.

REFERENCES:

[1] Foster, T. W., Jenkins, D. R., and Vickrey, D. "ADDITIONAL EVIDENCE ON THE INCREMENTAL INFORMATION CONTENT OF THE 10-K". Journal of Business Finance & Accounting. vol. 10. no. 1. pp. 57-66. Mar. 1983. 10.1111/j.1468-5957.1983.tb00412.x.
[2] F. Li, R. J. Lundholm and M. Minnis. "A Measure of Competition Based on 10-K Filings". Journal of Accounting Research. vol. 51. no. 2. pp. 399-436. Jan. 2013. 10.1111/j.1475-679x.2012.00472.x.
[3] C. S. A. Cheng, J. Fu, W. Huang and J. Jiao. "Strategic Use of Volume of Financial Items in 10-K Reports". Journal of Accounting, Auditing & Finance. vol. 38. no. 3. pp. 542-571. Feb. 2021. 10.1177/0148558x21989909.
[4] P. Lewis et al.. "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks". arXiv (Cornell University). May. 2020.
[5] S. Borgeaud et al.. "Improving language models by retrieving from trillions of tokens". arXiv (Cornell University). Dec. 2021. 10.48550/arxiv.2112.04426.
[6] B. Sarmah, T. Zhu, D. Mehta and S. Pasquali. "Towards reducing hallucination in extracting information from financial reports using Large Language Models". arXiv (Cornell University). Oct. 2023. 10.48550/arxiv.2310.10760.
[7] J. Z. Pan et al.. "Large Language Models and Knowledge Graphs: Opportunities and Challenges". arXiv (Cornell University). Aug. 2023. 10.48550/arxiv.2308.06374.
[8] "[2003.02320v2] Knowledge Graphs".
[9] R. Reinanda, E. Meij and M. D. Rijke. "Knowledge Graphs: An Information Retrieval Perspective". Foundations and Trends in Information Retrieval. vol. 14. no. 4. pp. 289-444. Jan. 2020. 10.1561/1500000063.
[10] "The Working Limitations of Large Language Models".
[11] SEC report - 10K filings https://www.sec.gov/ix?doc=/Archives/edgar/data/1160951/000119312523044867/d720670d10ka.htm
[12] SEC report - 10Q filings https://www.sec.gov/ix?doc=/Archives/edgar/data/1160951/000119312523044871/d775457d10qa.htm