GraphRAG Analysis for Financial Narrative Summarization and A Framework for Optimizing Domain Adaptation

Neelesh K Shukla, Prabhat Prabhakar, Sakthivel Thangaraj, Sandeep Singh, Weiyi Sun, C Prasanna Venkatesan, Viji Krishnamurthy
OCI Generative AI Services, Oracle Corporation
Correspondence: neelesh.kumar.shukla@oracle.com

Abstract

Large Language Models (LLMs) have shown promise in summarizing complex documents, but their limitations in handling lengthy documents and capturing global information hinder their performance in tasks like Query-Focused Summarization (QFS). To address these limitations, we explore GraphRAG, a retrieval-augmented generation approach that utilizes a globally summarized knowledge graph derived from an LLM. We apply GraphRAG to the Financial Narrative Summarization (FNS) dataset, which consists of lengthy financial reports. Our results show that a naive RAG approach outperforms GraphRAG in terms of comprehensiveness, directness, conciseness and completeness. However, we demonstrate that optimizing entity and relation extraction using an LLM as an optimizer can enhance GraphRAG's performance. Our study highlights the need for domain-specific optimization to improve GraphRAG's capabilities for summarization tasks in facts-heavy domains like finance. We propose an optimization framework that extends GraphRAG's original domain adaptation strategy by incorporating entity and relations optimization, leading to improved performance in capturing relevant entities and relationships. Our findings contribute to the development of more effective summarization models for complex documents in finance and other domains.

1 Introduction

Large Language Models (LLMs) have shown promise in analyzing complex documents and generating summaries, but they face significant challenges in summarizing lengthy documents due to restrictions on their context windows. The expansion of such windows may not be enough given that information can be "lost in the middle" of longer contexts (Liu et al., 2024). Retrieval-augmented generation (RAG) (Lewis et al., 2020) is a method that can overcome these limitations, but it struggles with capturing global information and addressing global queries, such as 'What are the main themes in the dataset?' This limitation, particularly its inability to effectively capture global information hinders its performance in tasks such as Query-Focused Summarization (QFS), where a broader understanding of the data is necessary (Peng et al., 2024). To address these limitations, an approach called GraphRAG (Edge et al., 2024) has been proposed, which utilizes a globally summarized knowledge graph derived from an LLM to unlock LLM discovery on narrative private data[1]. Building on this work, our research efforts focus on two key areas:

Our work is mainly directed towards two key areas:

- **Analysis of GraphRAG on financial narratives:** Previous research has explored the effectiveness of GraphRAG on datasets comprising podcast transcripts and news articles (Edge et al., 2024). We aim to broaden the scope by investigating the effectiveness of GraphRAG-based query-focused summarization in fact-rich domains like finance. Specifically, we apply GraphRAG to the Financial Narrative Summarization (FNS) shared task (Zavitsanos et al., 2023) which involves summarizing lengthy financial documents, such as annual reports around narrative sections. This makes this an ideal case study for the GraphRAG approach. The complexity of financial reports, characterized by technical terminology, numerical data, and domain-specific jargon, presents an ideal test case for GraphRAG's capabilities. To our knowledge, this study is the first to explore the application of GraphRAG to the FNS dataset, providing

[1] https://www.microsoft.com/en-us/research/blog/graphrag-unlocking-llm-discovery-on-narrative-private-data

Proceedings of the Joint Workshop of the 9th FinNLP, the 6th FNP, and the 1st LLMFinLegal, pages 23–34
January 19–20, 2025.

```
-Goal-
Given a text document that is potentially relevant to this activity and a list of entity types, identify all entities of those types from the text and all relationships among the
identified entities.

-Steps-
1. Identify all entities. For each identified entity, extract the following information:
- entity_name: Name of the entity, capitalized
- entity_type: One of the following types: {{entity_types}}
- entity_description: Comprehensive description of the entity's attributes and activities
Format each entity as ("entity"{tuple_delimiter}<entity_name>{tuple_delimiter}<entity_type>{tuple_delimiter}<entity_description>)

2. From the entities identified in step 1, identify all pairs of (source_entity, target_entity) that are *clearly related* to each other.
For each pair of related entities, extract the following information:
- source_entity: name of the source entity, as identified in step 1
- target_entity: name of the target entity, as identified in step 1
- relationship_description: explanation as to why you think the source entity and the target entity are related to each other
- relationship_strength: an integer score between 1 to 10, indicating strength of the relationship between the source entity and target entity
Format each relationship as
("relationship"{tuple_delimiter}<source_entity>{tuple_delimiter}<target_entity>{tuple_delimiter}<relationship_description>{tuple_delimiter}<relationship_strength>)

3. Return output in english as a single list of all the entities and relationships identified in steps 1 and 2. Use **{record_delimiter}** as the list delimiter.

4. If you have to translate into english, just translate the descriptions, nothing else!

5. When finished, output {completion_delimiter}.

######################
-Examples-
######################
Example 1:

Entity_types: {{entity_types}}
Text: {{example_text}}
################
Output:
("entity"{tuple_delimiter}THE RECORD HALL{tuple_delimiter}PRODUCT{tuple_delimiter}A Workspace product in Hatton Garden, featuring 89 units, roof terraces, Club Workspace, high-spec
meeting rooms, workshops for jewelry traders, and a new cafe partnership.}
{record_delimiter}
("entity"{tuple_delimiter}WORKSPACE{tuple_delimiter}ORGANIZATION,BRAND,WEBSITE{tuple_delimiter}Workspace is a company that offers a range of workspace products and services, with a
focus on the right market, properties, brand, customers, and people. Their website is www.workspace.co.uk.}
{record_delimiter}
#############################
-Real Data-
######################
Entity_types: {{entity_types}}
Text: {input_text}
######################
Output:
```

Figure 1: Example of Entity and Relationship Extraction Prompt

new insights into the model's performance in this challenging domain.

- **Optimizing domain adaptation:** We propose an optimization framework to enhance the performance of GraphRAG by incorporating entity and relation optimization. This framework ensures better alignment between ground-truth summaries and generated summaries with respect to an objective function, using an LLM as an optimizer.

2 Overview of Financial Narrative Summarization 2023 Dataset

The FNS 2023 task dataset[2] has been extracted from annual financial reports in PDF file format. The reports were written in English, Spanish, and Greek. For the dataset compilation, two to three people had to work for each language. For this work, we utilized English dataset which contains approximately 4,000 UK annual reports for firms listed on LSE, covering the period between 2002 and 2022 (El-Haj et al., 2014; El-Haj et al., 2022). In total, there are 4,013 annual reports divided into training, testing, and validation sets. Table 1 shows the dataset details.

Data Type	**Train**	**Validation**
Report Full-Text	3050	413
Gold Summaries	10.007	1383

Table 1: FNS 2023 Shared Task English Dataset

3 Background of GraphRAG and It's Domain Adaptation

3.1 Default GraphRAG

GraphRAG employs large language models (LLMs) to construct a detailed knowledge graph that captures entities and their relationships from a collection of text documents. This graph allows GraphRAG to utilize the semantic structure of the data to respond to complex queries, offering a broad contextual understanding. The process of creating this graph, known as indexing, involves guiding an LLM through the source content using domain-specific prompts. The LLM extracts relevant entities and relationships to form the graph. Key prompts used during the indexing process include: A) Entity and relationship extraction: Identifies entities and defines the relationships between them. B) Entity and relationship summarization: Merges instances of entities and relationships into a concise description. C) Community report generation: Provides summary reports for each community within the graph. These steps enable GraphRAG to efficiently organize and lever-

[2] http://wp.lancs.ac.uk/cfie/fns2023/

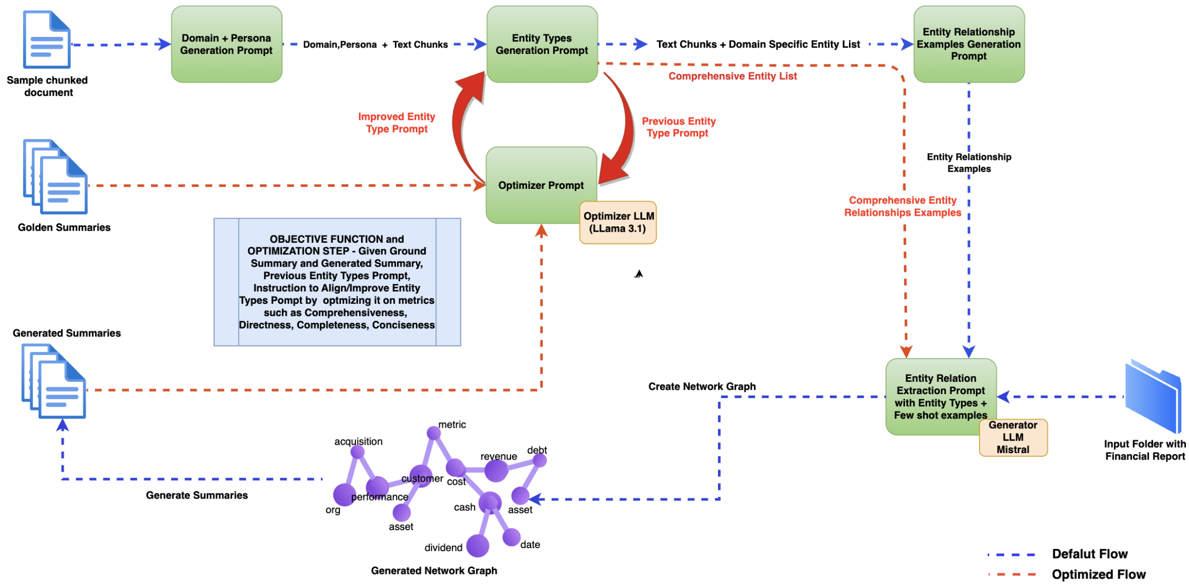

Figure 2: Domain Adaptation Flow: GraphRAG's Auto-Tuning Process (Blue) and Our Optimized Auto-Tuning Approach Utilizing LLM as an Optimizer (Red)

age the extracted knowledge for enhanced query responses[3].

We will compare Default GraphRAG with Naive RAG, a basic version that chunks the documents in fixed sizes and indexes, then uses cosine similarity to retrieve relevant chunks which combined with the original prompt to generate a response via an LLM.

3.2 GraphRAG's Approach for Domain Adaptation

Each domain possesses unique entity and relationship types, rendering manual prompt creation a time-intensive process. To address this, the GraphRAG team developed an automated tool for generating and refining domain-specific prompts efficiently. Consider the example of auto-tuning prompt for 'Entity and Relationship Extraction'. This prompt incorporates essential components: entity and relationship extraction instructions, few-shot examples, real data placeholders. An example is illustrated in Figure 1. The flow, illustrated in 'Blue' in Figure 2 demonstrates this approach. To begin, a sample of the source content is provided to the language model (LLM) to identify the domain and define a suitable persona. This persona is subsequently used in the Entity Type Generation Prompt to determine entity types relevant to the identified domain. Next, these domain-specific entity types are input into the 'Entity Relationship Example Generation' prompt to generate representative examples of relationships among entities within the domain. Finally, the extracted entity types and relationship examples are combined to construct a comprehensive 'Entity and Relationship Extraction' prompt. This prompt is employed by a Graph Generator LLM to extract entities and their relationships from any given text. The following are the entity types identified using this methodology. **Entity Types:** *[organization, market, location, financial metric, product, time]*

4 Optimizing Domain-Adaptation: Integrating LLM-as Optimizer and Ground Truth Summaries

The sole dependence on domain knowledge and persona-based methods for entity type identification is inadequate in capturing the dynamic nature of real-world data. While domain knowledge offers a baseline understanding and persona customization improves prompt design, these static strategies fall short in accommodating the intricate relationships and variations inherent in diverse datasets. To overcome this limitation, we propose an approach that integrates GraphRAG's domain adaptation with training data, leveraging ground truth summaries to enhance entity type recognition for enhanced domain-adaptation. Our proposed

[3]https://www.microsoft.com/en-us/research/blog/graphrag-unlocking-llm-discovery-on-narrative-private-data/

method employs Large Language Models (LLMs) as optimizer, framing the optimization task through natural language instructions. This enables the dynamic refinement of prompts, adapting them to the specific context and data nuances. By integrating this optimization process, we aim to achieve a more accurate and adaptable entity type identification.

Prompt for Optimizer LLM is designed in a three-step manner:

- **Comparison and Error Identification:** Similar to traditional machine learning, Optimizer LLM is asked compare the golden summary (actual output) with the generated summary (predicted output) to identify missing entities, relationships, or facts.
- **Objective Function and Analysis:** In this step, the objective of Optimizer LLM is to analyze the identified errors with focus on improving specific metrics, such as comprehensiveness, directness, completeness, and conciseness.
- **Instruction Generation:** Based on the analysis and insights from the previous step, Optimizer LLM is asked to generate instructions to enhance the entity relation extraction process in subsequent iterations, thereby leading to improved summary generation.

The optimizer prompt and objective function employed are as follows:

Optimizer Prompt and Objective Function:

- *Compare the GENERATED_SUMMARY with the GOLDEN_SUMMARY to assess how well the entities,and relationships were extracted and captured.*
- *Determine why these entities or relationships might have been missed.*
- *Make modifications to the ENTITY_TYPES_PROMPT that would improve the extraction of entities and relationships in the next iteration, based on the optimisation metrics provided (Comprehensiveness, Directness, Completeness, and Conciseness)*

As depicted in Figure 2, in 'Red', this iterative process begins with the Generator LLM producing summaries based on <Report Chunks, Ground Summaries>, utilizing original GraphRAG's domain adaptation technique described in Section 3.2. Subsequently, the Optimizer LLM, guided by an outlined optimizer prompt and objective function, evaluates the generated summaries against ground truth, iteratively generating recommendations to improve 'Entity Type Generation Prompt' to align generated entities and relationship with the specified objectives. This adaptive approach ensures a continuous improvement in entity type identification, leading to enhanced precision and an increased alignment with domain-specific objectives.

Below are the Optimizer LLM's final suggestions for refining the prompt, along with the identified entity types:

Recommendations for prompt refinement:
To improve the entity extraction prompt, I suggest modifying the ENTITY_TYPES_PROMPT

- *The user's task is to analyze the financial report and extract relevant entities and relationships.*
- *To include more specific entity types relevant to the task such as company, acquisition, financial metric, product, location, CEO, division, market, revenue, operating profit, cash generation, ROCE, health and safety, footprint, integration*
- *Change task in Real data section as: REAL DATA: Task: Analyze the financial report and extract relevant entities and relationships.*

Entity Types: *[organization, financial_report, metric, location, person, date, investment, revenue, customer, product, website, property, brand, safety, certification, acquisition, debt, employee, factory, construction, asset, cost, strategy, cash, dividend, drilling, committee, principle, environmental_impact, growth, appointment, performance, acquisition, sales]*

5 Experimental Setup

Instead of summarizing the complete report, the FNS task requires locating key narrative sections found in the annual reports and generate a single structured summary for them in not more than 1000 words (Figure 3). We utilized DiMSum (Shukla et al., 2022) for narrative section identification and extraction . This system was the top performer in the FNS 2023 task (Zavitsanos et al., 2023).

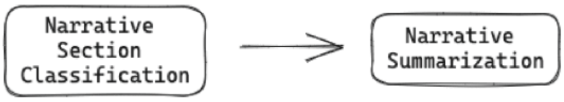

Figure 3: Two step summarization

5.1 Query

To evaluate the effectiveness of RAG systems on FNS task, we formed the query that convey the task requirement and only a high-level understanding of dataset contents.

Query: *Please extract narrative summary of [COMPANY_NAME]'s annual financial report in not more than 1000 words.*

5.2 Evaluation Metrics

Large Language Models (LLMs) have been shown to be effective in evaluating natural language generation, achieving results comparable to human judgments (Wang et al., 2023; Zheng et al., 2024). To assess the quality of generated text, we employed four metrics that utilize LLMs as evaluators.

For direct comparison, we adapted two metrics from GraphRAG for FNS task:

- *Comprehensiveness:* Does the system summary adequately cover all relevant details found in the human summaries? Evaluate how well it captures the breadth and depth of key information.

- *Directness:* How concise and straightforward is the system summary? Assess the extent to which it clearly and effectively distills the essential points from the human summaries without unnecessary complexity.

Additionally, we used FineSurE (Song et al., 2024), a fine-grained summarization evaluation approach that leverages LLMs to evaluate summary quality at a detailed level. This method identifies key facts utilizing LLMs, which are concise sentences conveying a single piece of information (Bhandari et al., 2020), and evaluates summaries based on two metrics:

- *Conciseness:* Avoiding unnecessary details. Interpreted as precision of Key Facts.

- *Completeness:* Encompassing the majority of key facts in the summary. Interpreted as Recall of Key Facts.

5.3 Configurations

GraphRAG is designed to use Microsoft Supported LLMs and Embedding (OpenAI Models). In our experiments, we employ Ollama's Mistral-7B LLM[4] and Nomic-Embed-Text[5] embedding model due to limited access to Microsoft's models. Consistent settings is applied across all experiments: chunk size (1200), overlap (100), and summary length (1000). ChromaDB is used as vector store in NaiveRAG. For domain-adaptation, taking advantage of Llama3.1-405B's larger context window and expanded parameter set, it is used for generating various prompts and LLM as an optimizer, as depicted in Figure 2. System performance is evaluated by a Judge LLM, Cohere Command R+[6], which is a separate LLM from the generator and optimizer LLM.

6 Results and Analysis

Our results (Table 2) reveal that the Naive RAG approach surpasses the GraphRAG on FNS. Appendix A, contains examples of summaries. The key takeaways from the analysis are summarized below.

- Comprehensiveness: Naive RAG provides more comprehensive summaries, capturing key aspects of the financial reports, including financial highlights, performance, strategy, and market trends In contrast Graph RAG's focus on broader themes such as role of employees and the management development program, but omitting detailed financial and strategic insights which may limit its usefulness for stakeholders seeking detailed financial information.

- Directness: Naive RAG exhibits a higher degree of directness, maintaining a tight alignment with the source material and concentrating on key financial metrics, financial performance, strategic initiatives, and outlook. In contrast, Graph RAG tends to deviate from the main theme, emphasizing peripheral aspects such as employee contributions, community dynamics, and external events, rather than providing a straightforward account of financial performance and strategic initiatives, due to

[4]https://ollama.com/library/mistral
[5]https://ollama.com/library/nomic-embed-text
[6]https://docs.cohere.com/v2/docs/command-r-plus

Approach	Comprehen-siveness	Directness	Complete-ness	Concise-ness
Default GraphRAG	57.66	67.48	5.99	18.57
Naive RAG	79.81	79.79	27.18	49.53

Table 2: Comparison of NaiveRAG vs GraphRAG on Validation Dataset

Approach	Comprehen-siveness	Directness	Complete-ness	Concise-ness
Default GraphRAG	57.66	67.48	5.99	18.57
GraphRAG's Domain Adaptation	67.81	79.69	10.43	26.06
Optimized Domain Adaptation	**75.45**	**83.17**	10.04	24

Table 3: GraphRAG Domain Adaptation Results: Comparison of GraphRAG's Domain Adaptation vs. Our Optimized Domain Adaptation on Validation Dataset.

its prioritization of entity relationships that can introduce tangential information.

- Completeness: The Naive RAG approach achieves a high degree of completeness, successfully extracting key financial metrics, strategic information, and important business details from the input text. In contrast, the Graph RAG approach falls short, frequently omitting crucial metrics and details that are present in the reference summaries. This disparity in performance is attributed to the limitations of the graph-based approach, specifically its structure and entity list, which hinder its capacity to thoroughly retrieve relevant information, ultimately leading to less comprehensive summaries.

- Conciseness: Our evaluation reveals that Naive RAG generates concise summaries, effectively balancing brevity and informativeness by focusing on key financial figures and insights without unnecessary elaborations. In contrast, Graph RAG sometimes includes irrelevant or overly abstract information, reducing its precision and conciseness. Specifically, it occasionally introduces extraneous concepts and details not directly related to the main topic of the financial report, detracting from the summary's focus.

Our optimized domain adaptation approach enhanced Graph RAG's ability to generate more accurate and detailed summaries by embedding enriched entity relationships as context. As shown in Appendix A (Table 6, 7), summaries contains relevant entities like revenue, net income, and operating expenses, resulting in a more comprehensive summary. The broader entity list improved coverage of key financial and operational concepts, while entity relationships provided deeper insights into interconnected financial details. The expanded graph structure included both strategic and granular financial metrics. Our results (Table 3) demonstrate improvements in comprehensiveness and directness, reflecting the better capture of relevant entities.

Despite the optimization, Naive RAG still outperforms GraphRAG (Table 2,3). The Naive RAG technique achieves a highly relevant summary by directly integrating information from the source document, effectively capturing key financial metrics and contextual elements. Unlike GraphRAG, Naive RAG successfully identifies critical aspects such as acquisition targets, executive leadership changes, and the impact of external factors like Brexit. This direct integration results in summaries that align closely with ground-truth references.

In contrast, while entity recognition is improved, GraphRAG's ability to extract all relevant entities and establish detailed relationships remains limited. This constraint hampers its capacity to construct a comprehensive knowledge graph. By prioritizing relational and community-level summarization, GraphRAG often sacrifices critical details, leading to summaries that are high-level and less informative. For example, it mentions growth trends without providing comparative figures and references acquisitions without specifying details. Furthermore, it omits external contextual factors, such as Brexit, which are essential for a nuanced analysis.

7 Conclusion and Future Work

In conclusion, our study reveals that Naive RAG outperforms Graph RAG in extracting actionable insights from financial metrics, strategies, and market trends, as the latter's reliance on identified entities can lead to off-topic deviations. Nevertheless, Graph RAG shows potential by harnessing the power of training data and Large Language Models (LLMs) as optimizers, which successfully align entities and relationships to produce relevant objective summaries. Notably, our findings suggest that the incorporation of an entity-relationship graph does not necessarily guarantee superior content summaries. Future studies would benefit from incorporating metrics to assess the accuracy of extracted entities and relationships, offering insights into GraphRAG's potential in this domain. Exploring methods to construct graphs based on thematic elements, such as revenue performance variations, and generating subgraphs for each theme, could further enhance GraphRAG's capabilities. Thematic summaries can be employed to identify relevant entities and relationships tied to specific themes, improving the graph's alignment with the narrative content. Our findings underscore the importance of domain-specific optimization to refine GraphRAG for summarization tasks in data-intensive sectors like finance.

Ethical Considerations

This research emphasizes ethical considerations by basing all claims on experimental results, ensuring transparent documentation of methodologies, and sourcing datasets ethically with the necessary permissions.

References

Manik Bhandari, Pranav Narayan Gour, Atabak Ashfaq, Pengfei Liu, and Graham Neubig. 2020. Re-evaluating evaluation in text summarization. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 9347–9359, Online. Association for Computational Linguistics.

Darren Edge, Ha Trinh, Newman Cheng, Joshua Bradley, Alex Chao, Apurva Mody, Steven Truitt, and Jonathan Larson. 2024. From local to global: A graph rag approach to query-focused summarization.

Mahmoud El-Haj, Paul Rayson, Paulo Alves, Carlos Herrero-Zorita, and Steven Young. *Multilingual Financial Narrative Processing: Analyzing Annual Reports in English, Spanish, and Portuguese*, chapter Chapter 13.

Mahmoud El-Haj, Paul Rayson, Steve Young, and Martin Walker. 2014. Detecting document structure in a very large corpus of UK financial reports. In *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*, pages 1335–1338, Reykjavik, Iceland. European Language Resources Association (ELRA).

Mahmoud El-Haj, Paul Rayson, and Nadhem Zmandar, editors. 2022. *Proceedings of the 4th Financial Narrative Processing Workshop @LREC2022*. European Language Resources Association, Marseille, France.

Patrick Lewis, Ethan Perez, Aleksandra Piktus, Fabio Petroni, Vladimir Karpukhin, Naman Goyal, Heinrich Küttler, Mike Lewis, Wen-tau Yih, Tim Rocktäschel, Sebastian Riedel, and Douwe Kiela. 2020. Retrieval-augmented generation for knowledge-intensive nlp tasks. In *Proceedings of the 34th International Conference on Neural Information Processing Systems*, NIPS '20, Red Hook, NY, USA. Curran Associates Inc.

Nelson F. Liu, Kevin Lin, John Hewitt, Ashwin Paranjape, Michele Bevilacqua, Fabio Petroni, and Percy Liang. 2024. Lost in the middle: How language models use long contexts. *Transactions of the Association for Computational Linguistics*, 12:157–173.

Boci Peng, Yun Zhu, Yongchao Liu, Xiaohe Bo, Haizhou Shi, Chuntao Hong, Yan Zhang, and Siliang Tang. 2024. Graph retrieval-augmented generation: A survey. *Preprint*, arXiv:2408.08921.

Neelesh Shukla, Amit Vaid, Raghu Katikeri, Sangeeth Keeriyadath, and Msp Raja. 2022. DiMSum: Distributed and multilingual summarization of financial narratives. In *Proceedings of the 4th Financial Narrative Processing Workshop @LREC2022*, pages 65–72, Marseille, France. European Language Resources Association.

Hwanjun Song, Hang Su, Igor Shalyminov, Jason Cai, and Saab Mansour. 2024. FineSurE: Fine-grained summarization evaluation using LLMs. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 906–922, Bangkok, Thailand. Association for Computational Linguistics.

Jiaan Wang, Yunlong Liang, Fandong Meng, Zengkui Sun, Haoxiang Shi, Zhixu Li, Jinan Xu, Jianfeng Qu, and Jie Zhou. 2023. Is ChatGPT a good NLG evaluator? a preliminary study. In *Proceedings of the 4th New Frontiers in Summarization Workshop*, pages 1–11, Singapore. Association for Computational Linguistics.

Elias Zavitsanos, Aris Kosmopoulos, George Giannakopoulos, Marina Litvak, Blanca Carbajo-Coronado, Antonio Moreno-Sandoval, and Mo El-Haj. 2023. The financial narrative summarisation shared task (fns 2023). In *2023 IEEE International Conference on Big Data (BigData)*, pages 2890–2896.

Lianmin Zheng, Wei-Lin Chiang, Ying Sheng, Siyuan Zhuang, Zhanghao Wu, Yonghao Zhuang, Zi Lin, Zhuohan Li, Dacheng Li, Eric P. Xing, Hao Zhang, Joseph E. Gonzalez, and Ion Stoica. 2024. Judging llm-as-a-judge with mt-bench and chatbot arena. In *Proceedings of the 37th International Conference on Neural Information Processing Systems*, NIPS '23, Red Hook, NY, USA. Curran Associates Inc.

A Appendix: Examples of Summaries generated

This appendix includes example summaries for Vodafone Group PLC and Volution Group PLC, generated using two different methods: Naive RAG vs GraphRAG (Tables 4 and 5), as well as Auto-Tuning vs our optimized Auto-Tuning approach (Tables 6 and 7).

Naive RAG	**Default GraphRAG**
Vodafone Group Plc, in its Annual Report for the year 2022, highlights its commitment to enabling an inclusive and sustainable digital society while ensuring responsible business practices. The report covers strategic, operational, and financial progress. In terms of financial performance, the Group's revenue stood at £45,580 million in 2022, representing a growth from £43,809 million in 2021 and a slight decrease from £44,974 million in 2020. The Group's service revenue amounted to £38,203 million in 2022, up from £37,141 million in 2021 but slightly lower than the £37,871 million recorded in 2020. Employee engagement was another area of focus for Vodafone. The average number of employees and contractors remained relatively stable at 104 thousand in both 2021 and 2022, compared to 103 thousand in 2020. Employee engagement index also saw a slight increase from 73% in 2021 to 74% in 2022. Vodafone prides itself on its responsible business practices and operating lawfully with integrity. The company's progress towards its purpose is reflected in these key performance indicators. The Annual Report also includes references to the Group's website for additional information, such as videos, ESG Addendum, and TCFD report. Readers are advised that online content is for convenience only and does not form part of this annual report. The document provided is the UK version of the Annual Report, with the US Securities and Exchange Commission (SEC) filing to follow at a later date. Some of the highlights from Vodafone's FY22 include: - Strong focus on digital society and responsible business practices - Financial progress: Group revenue increased from £43.8 billion in 2021 to £45.6 billion in 2022, with a slight decrease in service revenue from £37.9 billion in 2020 to £38.2 billion in 2022 - Employee engagement index rose slightly from 73% in 2021 to 74% in 2022 - Commitment to sustainability and ethical business practices The Annual Report emphasizes that the company's success is measured through its key performance indicators, which reflect both financial and non-financial aspects of its business. These indicators help underpin the delivery of Vodafone's purpose and ensure sustainable, long-term growth.	Vodafone Group Plc, a key player within the Europe TV Subscribers community, has shown a significant financial standing with a total revenue of €46,123 million in the year under review [Data: Reports (1)]. This substantial income suggests a strong position within this market. The report also reveals that Vodafone Group Plc's net profit for the year was €4,286 million, indicating a profitable operation within the Europe TV Subscribers community [Data: Reports (1)]. This profitability underscores the company's ability to generate earnings and reinvest in its services. At the end of the year under review, Vodafone Group Plc's total assets were €84,672 million, highlighting its substantial financial resources within this community [Data: Reports (1)]. This asset base provides a strong foundation for future growth and expansion. While the annual financial report does not provide specific information about Vodafone Group Plc's subscriber base within the Europe TV Subscribers community, it is known that this community has a large number of subscribers [Data: Entities]. This suggests potential opportunities for growth and expansion for Vodafone Group Plc within this market. In conclusion, Vodafone Group Plc's annual financial report presents a positive outlook for the company within the Europe TV Subscribers community. With its strong financial standing, profitable operation, substantial resources, and potential for growth, Vodafone Group Plc is well-positioned to capitalize on opportunities in this market.

Table 4: Comparision of Sumamries Genertated by NaiveRAG vs Default GraphRAG for Vodafone Group PLC.

Naive RAG	**Default GraphRAG**
Volution Group plc, in its Annual Report for the year ended July 31, 2017, reported strong financial results despite the uncertainties brought about by the UK's vote to leave the European Union and the devaluation of Sterling against major currencies. The company's strategy revolves around building a more valuable Volution Group through brand portfolio expansion and market and geographical diversification. The financial highlights revealed a 19.8% increase in revenue, with figures reaching £185.1 million in 2017 compared to £154.5 million in 2016. Adjusted operating profit grew by 9.6% to £35.6 million (19.3% of revenue), and there was a significant increase in adjusted EPS, with basic EPS standing at 35.9p and diluted EPS at 32.5p. The net debt stood at £37.0 million, an increase from the previous year's £21.2 million. Reported profit before tax was £17.9m with a loss of (4.2)p per share. Volution Group reported strong cash generation and continued revenue and underlying operating profit growth for the sixth consecutive year. The company also uses alternative performance measures to assess the business's underlying performance, including adjusted operating profit, adjusted profit before tax, adjusted basic and diluted EPS, and adjusted operating cash flow. Chairman Peter Hill expressed his satisfaction with the progress made during the year and the company's confidence in its long-term prospects despite Brexit uncertainties. The Group continued to make strategic moves by making selective value-adding acquisitions and expansions, all while focusing on growth opportunities in both established and emerging markets. In conclusion, Volution Group plc reported a successful financial year with revenue and underlying operating profit growth, significant cash generation, and strong strategic progress, despite the economic uncertainties brought about by Brexit. The company remains committed to its strategy of building a more valuable Volution Group through brand portfolio expansion and market and geographical diversification.	Volution Group PLC, a company that has integrated two new businesses, attributes its success to the significant contributions made by its employees [Data: Reports (0)]. The central entity within the Verdant Oasis Plaza community, Volution Group PLC, values its Management Development Programme which plays a pivotal role in the integration of new acquisitions and the formation of the overall Group culture [Data: Reports (0)]. The employees of Volution Group PLC are not only integral to its success but also significantly contribute to the organization's performance [Data: Reports (0)]. The dynamics of the Verdant Oasis Plaza community are crucial in understanding the relationship between the plaza and any events that may take place there, such as the Unity March, which could potentially pose a threat [Data: Reports (0)]. Currently, Verdant Oasis Plaza is experiencing media attention due to the Unity March taking place within the plaza. This increased exposure could amplify its impact on the community [Data: Reports (0)]. It's essential to consider the nature of the Unity March and the reactions it provokes when analyzing the relationship between the march and the plaza. In conclusion, Volution Group PLC's success is largely due to the contributions made by its employees, and its Management Development Programme plays a crucial role in integrating new acquisitions and shaping the overall Group culture. The dynamics of the Verdant Oasis Plaza community are significant when understanding events that take place within the plaza, such as the Unity March. The increased media attention on the Unity March could potentially impact the community. For more detailed information, please refer to the relevant data records (0, +more)

Table 5: Comparision of Sumamries Genertated by NaiveRAG vs Default GraphRAG for Volution Group PLC.

GraphRAG's Domain Adaptation	**Optimized Domain Adaptation**
Vodafone Group Plc, a leading telecommunications company, reported a total revenue of £46.2 billion for the fiscal year [Data: Reports (1)]. This significant financial performance indicates a strong position in the market. However, it is important to note that the company's net debt stood at £38.7 billion, demonstrating a substantial debt level that could impact future financial decisions [Data: Reports (2)]. The report highlighted a focus on cost optimization and network investments, aiming to improve profitability and customer experience [Data: Reports (4)]. Despite these efforts, Vodafone Group Plc experienced a decline in EBITDA margin to 31.4%, suggesting operational efficiency challenges [Data: Reports (3)]. On a positive note, Vodafone Group Plc's free cash flow improved to £7.2 billion, signaling a positive cash position for the company [Data: Reports (5)]. This improvement in cash flow could provide the necessary resources for the company to address its debt and operational efficiency issues. In conclusion, while Vodafone Group Plc demonstrated a strong financial performance with a significant revenue of £46.2 billion, it also faces challenges in terms of operational efficiency and a substantial debt level. The company's focus on cost optimization and network investments, as well as the improvement in free cash flow, could potentially help address these issues. Further analysis of additional reports [Data: Reports (+more)] may provide more insights into the company's future financial decisions and performance.	Vodafone Group Plc, a significant player in the IoT sector, reported approximately 150.1 million IoT SIM connections [Data: IOT SIM CONNECTIONS MILLION (top 5 relevant record ids)], indicating a growing interest in IoT services and potential for further expansion. This figure suggests that Vodafone is well-positioned to capitalize on the increasing demand for IoT solutions. The total revenue generated by Vodafone Group Plc stands at a substantial 24.7 billion Kenyan shillings [Data: REVENUE (top 5 relevant record ids)], indicating a strong financial performance and the ability to invest in future growth initiatives. This revenue figure underscores the company's robust market position and its capacity to drive innovation. However, it is important to note that while Vodafone Group Plc is profitable, there is room for improvement in managing costs. The current net income stands at 6.8 billion Kenyan shillings, while operating expenses amount to 17.9 billion Kenyan shillings [Data: NET INCOME (top 5 relevant record ids), OPERATING EXPENSES (top 5 relevant record ids)]. This implies that the company could optimize its operations to increase profitability. The IoT Service Provider Community report, while not explicitly naming Vodafone Group Plc, suggests that the company's IoT operations are part of a broader context [Data: Reports (other than IOT SIM CONNECTIONS MILLION, REVENUE, NET INCOME, OPERATING EXPENSES)]. This report provides insights into the performance and growth trends within the IoT service provider community, which could be beneficial for Vodafone in strategizing its future moves. Regarding ownership structure and allegations of wrongdoing, these details are not explicitly provided in the IoT Service Provider Community report [Data: Reports (other than IOT SIM CONNECTIONS MILLION, REVENUE, NET INCOME, OPERATING EXPENSES)]. For such information, it is recommended to consult other relevant reports [Data: Reports (2, 7, 34, 46, 64, +more)].

Table 6: Comparison of Summaries Generated by GraphRAG's Domain Adaptation vs. Our Optimized Domain Adaptation for Vodafone Group PLC.

GraphRAG's Domain Adaptation	**Optimized Domain Adaptation**
Volution Group PLC, a company that has integrated two new businesses, attributes its success to the significant contributions made by its employees [Data: Reports (0)]. The central entity within the Verdant Oasis Plaza community, Volution Group PLC, values its Management Development Programme which plays a pivotal role in the integration of new acquisitions and the formation of the overall Group culture [Data: Reports (0)]. The employees of Volution Group PLC are not only integral to its success but also significantly contribute to the organization's performance [Data: Reports (0)]. The dynamics of the Verdant Oasis Plaza community are crucial in understanding the relationship between the plaza and any events that may take place there, such as the Unity March, which could potentially pose a threat [Data: Reports (0)]. Currently, Verdant Oasis Plaza is experiencing media attention due to the Unity March taking place within the plaza. This increased exposure could amplify its impact on the community [Data: Reports (0)]. It's essential to consider the nature of the Unity March and the reactions it provokes when analyzing the relationship between the march and the plaza. In conclusion, Volution Group PLC's success is largely due to the contributions made by its employees, and its Management Development Programme plays a crucial role in integrating new acquisitions and shaping the overall Group culture. The dynamics of the Verdant Oasis Plaza community are significant when understanding events that take place within the plaza, such as the Unity March. The increased media attention on the Unity March could potentially impact the community. For more detailed information, please refer to the relevant data records (0, +more).	Volution Group plc, a leading international ventilation technology group, has announced its annual financial results for the year ended December 31, 2021. The report highlights a strong performance across various key areas, demonstrating resilience and growth despite the challenging global economic climate. Revenue for the year increased by 14.6% to £257.8 million (2020: £223.9 million), driven primarily by organic growth and strategic acquisitions. The Group's operating profit rose significantly, up 37.6% to £50.2 million (2020: £36.6 million). This improvement was underpinned by a focus on operational efficiency, cost management, and the successful integration of acquired businesses. The Group's adjusted earnings before interest, taxes, depreciation, and amortization (EBITDA) increased by 27.4% to £59.8 million (2020: £46.3 million), reflecting improved profitability across the business. The reported profit before tax was £41.2 million, compared to £28.8 million in 2020, representing a growth of 43.7%. The Group's net cash position at year-end stood at £56.9 million (2020: £34.4 million), demonstrating strong cash generation and financial discipline. This robust cash position will support the Group's ongoing investment in innovation, strategic acquisitions, and organic growth opportunities. The Board proposes a final dividend of 17.5 pence per share (2020: 13.0 pence), bringing the total dividend for the year to 24.0 pence per share, an increase of 89.3% compared to the previous year. This reflects the Group's strong financial performance and commitment to returning value to shareholders. Looking ahead, Volution Group remains confident in its strategic direction and market opportunities. The Group will continue to focus on innovation, operational efficiency, and strategic acquisitions to drive growth and create long-term shareholder value. The management team is excited about the prospects for the future and looks forward to navigating the challenges and opportunities that lie ahead.

Table 7: Comparison of Summaries Generated by GraphRAG's Domain Adaptation vs. Our Optimized Domain Adaptation for Volution Group PLC.