**Enhancing SOX ITGC audit compliance with LLM-Graph RAG based Chatbot**

**What is the idea? Explain in detail:**

SOX compliance refers to annual audits that take place within public companies based in US. But every year this process takes humongous manual efforts. According to Public Company Accounting Oversight Board (PCAOB)’s (nodal agency responsible for SOX compliance) report 2022, there were 34% audit quality deficiencies.

By providing SOX auditors with a reliable document-analysis LLM chatbot, we can streamline data analysis, documentation review, and anomaly detection in ITGC audits. These chatbots can quickly analyse large amounts of data, spot patterns or discrepancies that may signal control issues, and can assist in generating compliance reports more efficiently. This approach not only helps manage the workload but also improves accuracy and consistency.

**Problem Statement with brief background:**

Current RAG (Retrieval-Augmented Generation) chatbots works by first searching through a large dataset to find relevant information using semantic search. Then, it uses a language model to create a response based on that information. However, it often sufferes from keeping track of context and how different pieces of information relate to each other, which can result in **less accurate and less coherent answers.**

GraphRAG (Graph-based Retrieval-Augmented Generation) uses graph-based data structures along with advanced retrieval methods to improve the generation of accurate and relevant information. This makes it good at keeping track of relationships and connections in data, making it well-suited for complex tasks like SOX audits.

**Solution with implementation approach in brief:**

In this solution we are going to use Knowledge Graphs. It works by representing information as a network of interconnected entities and their relationships. Each entity (node) and relationship (edge) are labelled, enabling efficient querying and discovery of complex, structured data. Sample Knowledge graph is show in [Figure 1].

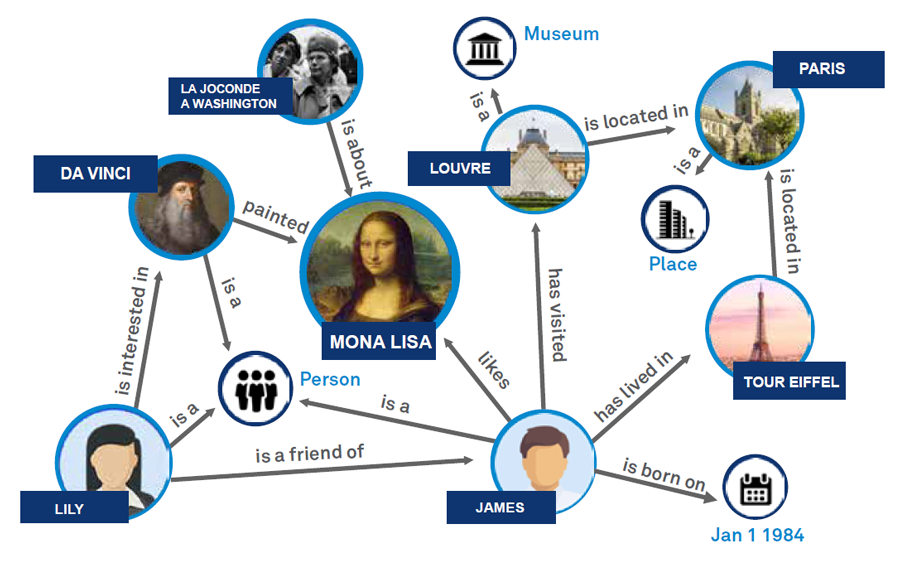


Figure 1 Sample Knowledge graph

First, we convert corporate data into knowledge graphs. Then, we classify these graphs into different communities (categories like Paris, Monalisa, James, etc.) and create community summaries. These summaries provide the LLM with a comprehensive overview by summarizing information from various sources, offering a well-rounded perspective on the topic.

When a user query is made, the relevant community summaries are combined with the query to generate a detailed report. This architecture efficiently answers user queries from complex, interconnected data. Unlike traditional RAG systems, GraphRAG can extract intricate relationships and insights from the data. High level flow diagram is given in [Figure 2].

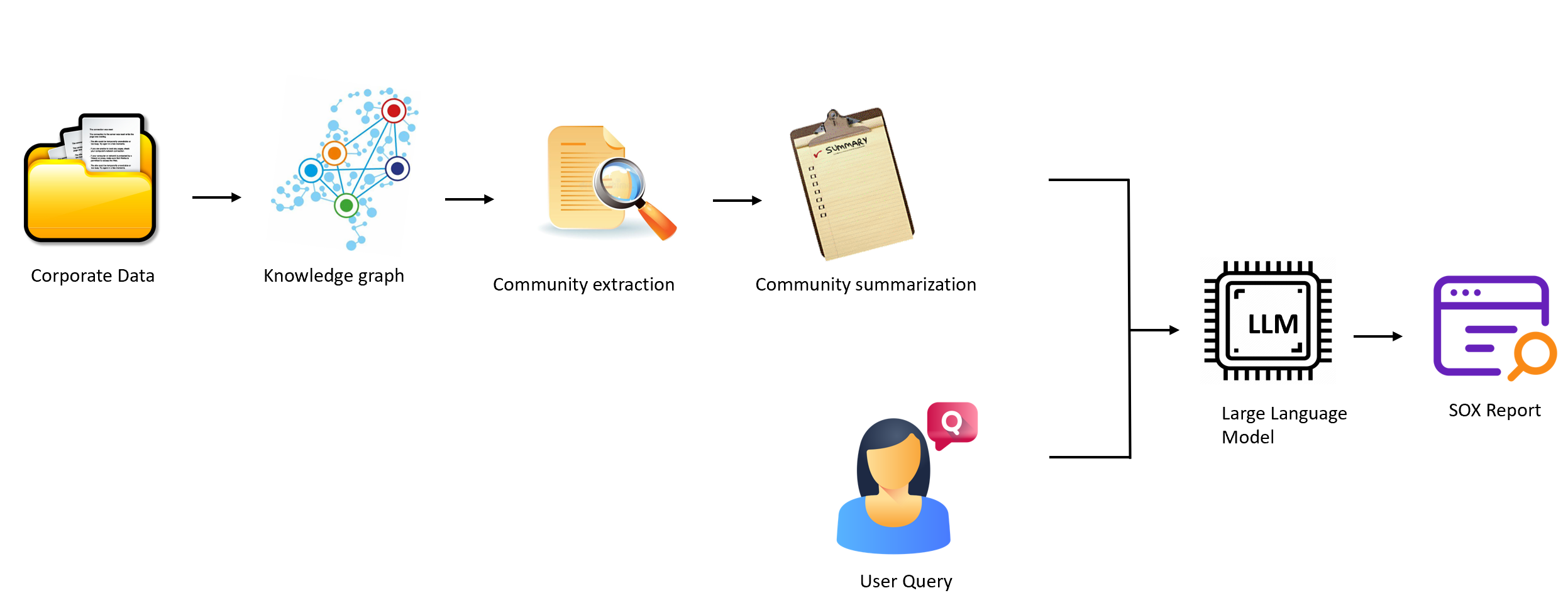


Figure 2 Solution flow diagram

**Value delivery with quantifiable impact/KPI:**

The below [Tabel 1] list the potential KPI’s that can be improved by incorporating the Knowledge graph based GenAI chatbot in audit process.

|  |  |  |
| --- | --- | --- |
| Sno | KPI | Comment |
| 1 | Compliance score | Can improve accuracy and consistency in SOX audits. |
| 2 | Project timelines | Can quickly identifying compliance issues, and can provide real-time insights |
| 3 | Employee productivity | Can reduce manual workload by automating repetitive tasks, allowing auditors to focus on higher-value activities. |
| 4 | IT Return on Investment (ROI) | Can reduce cost and saves time. |

Table 1: KPI's improved

# **References:**

1. **Public Company Accounting Oversight Board (PCAOB).** *Staff Update and Preview of 2022 Inspection Observations.* s.l. : PCAOB, July 2023.

2. *From Local to Global: A Graph RAG Approach to Query-Focused Summarization.* **Larson, Darren Edge and Ha Trinh and Newman Cheng and Joshua Bradley and Alex Chao and Apurva Mody and Steven Truitt and Jonathan.** 2404.16130, 2024, arXiv.