Scientific Research Publishing

Open Journal of Business and Management, 2026, 14(1), 628-653
https://www.scirp.org/journal/ojbm
ISSN Online: 2329-3292
ISSN Print: 2329-3284

# Knowledge Graph Application in KM for Accounting and Supply Chain in SMEs

**Yaxin Zheng**

Grant Thornton China, Beijing, China
Email: evazheng2003@qq.com

**How to cite this paper:** Zheng, Y. X. (2026). Knowledge Graph Application in KM for Accounting and Supply Chain in SMEs. *Open Journal of Business and Management, 14,* 628-653.
https://doi.org/10.4236/ojbm.2026.141038

**Received:** December 3, 2025
**Accepted:** January 18, 2026
**Published:** January 21, 2026



## Abstract

In today's data-driven business environment, small and medium-sized enterprises (SMEs) struggle to implement effective knowledge management (KM) due to limited financial, technical, and human resources. This study proposes an intelligent KM framework that integrates large language models (LLMs), that is, BERT-based models and knowledge graphs (KGs) to support accounting and supply chain decision-making in SMEs. Using a medical device manufacturer as a case enterprise, unstructured texts from accounting vouchers, ERP operation logs and supply chain records are collected, cleaned, anonymised, and manually annotated with entities and relations via Doccano. An end-to-end architecture based on BERT is adopted: for named entity recognition (NER) two models (BERT + CRF and BERT + BiLSTM + CRF) are trained, while for relation extraction (RE) two models (BERT and BERT + BiLSTM) are compared. Experimental results show that BERT + BiLSTM + CRF achieves the best NER performance (Precision 0.73, Recall 0.76, F1-score 0.74), and BERT attains superior RE performance (Precision 0.69, Recall 0.80, F1-score 0.74). The optimal models are then used to automatically construct RDF-like entity-relation-entity triples, which are stored and visualized in the Neo4j Aura cloud graph database via Py2neo. The resulting enterprise KG supports semantic querying, knowledge discovery, and interactive Q&A for accounting and supply chain tasks, and can be further leveraged for personalized training and internal audit support. The findings demonstrate that an LLM-enabled KG approach can provide a cost-effective, scalable KM solution for SMEs, while also highlighting challenges related to data quality, computational resources, and ongoing KG maintenance.

## Keywords

Knowledge Graph, Large Language Model, BERT, Knowledge Management, Small and Medium-Sized Enterprises, Accounting, Supply Chain Management

## 1. Introduction

In today's data-driven decision-making era, efficient knowledge management (KM) has become a key factor for enterprises to maintain competitiveness, particularly for small and medium-sized enterprises (SMEs) (SIMA et al., 2022; Durst, Edvardsson, & Foli, 2023). How to effectively manage knowledge assets and enhance decision intelligence under limited financial and human resources has become a critical issue for sustaining competitive advantages. Leveraging its strengths in semantic integration, intelligent retrieval, and knowledge representation, KM is widely applied in key business scenarios such as accounting and supply chain management for large enterprises, significantly improving the efficiency of information acquisition and utilization (Puspita & Pramono, 2019; Eslami et al., 2023; Dehshiri et al., 2024; Theuma, Hayak, & Khdour, 2025). However, constrained by limited funding, technology, and human resources, SMEs face practical challenges in constructing and applying knowledge graphs, including high costs, complex construction, and intricate maintenance (Koliby et al., 2025).

The rapid advancement of large language models (LLMs) in recent years has opened new possibilities for building intelligent knowledge graphs (KG) (Hu, Liu, & Dai, 2024). LLMs, with their robust natural language understanding and generation capabilities, can efficiently extract and organize unstructured data, thereby enhancing the automation and accuracy of KG construction. Current research primarily focuses on KM systems in large enterprises or specific industries (e.g., healthcare, legal, finance), emphasizing the impact of KM on performance improvement and competitive edge (Plessis & Toit, 2006; SIMA et al., 2022; Li et al., 2024; Pisoni, Molnar, & Tarci, 2024; Mukarram et al., 2025; Pereira & Fernandes, 2025).

However, research on knowledge collaboration in two core enterprise domains—supply chain and accounting—remains insufficient, lacking a systematic theoretical and methodological framework. Existing solutions face limitations in feasibility and portability due to SMEs' often incomplete data governance systems and inadequate digital infrastructure, making direct application challenging. This study therefore focuses on exploring how to leverage the natural language processing (NLP) capabilities of large language models and the structured representation advantages of knowledge graphs to develop an intelligent KM framework tailored for SMEs. The research selects accounting and supply chain as KG construction and validation scenarios, utilizing text information extraction, relationship identification, and graph visualization technologies to achieve intelligent transformation of knowledge from fragmented texts into systematic knowledge bases.

## 2. Literature Review

### 2.1. Analysis of the Effect of KG on LLM

In corporate financial management, KM has long been a key strategy to enhance decision-making efficiency and employee competency. However, enterprises face three major challenges: fragmented information that is difficult to integrate and standardize; difficulties in retrieving accounting knowledge due to limitations in search algorithms, inconsistent data formats, and information fragmentation, which may result in redundant, error-prone, or misleading results; and the lack of effective internal KM mechanisms, making new employee training particularly challenging.

To address these issues, Chen & Sun (2024) conducted a study on the tobacco industry, employing NLP technology to develop a KM platform tailored for corporate accounting. Their research utilizes Neo4j's high-performance graph database technology to construct an accounting KG, integrating fragmented information from multiple channels into a unified knowledge system that achieves comprehensive coverage of accounting knowledge. The study demonstrates that the application of this KG achieves over 80% accuracy, effectively supporting financial and tax professionals in acquiring relevant knowledge while enhancing work efficiency and quality. This finding validate the effectiveness of LLM in building KM platforms based on KG, offering valuable insights for SMEs in knowledge management. However, it possesses more resources to implement and maintain KM systems as the research focuses on large enterprises with greater technical, financial, and human resource advantages compared to SMEs. SMEs must consider factors such as cost, time investment, and long-term maintenance when establishing KM platforms based on KG.

Zhao et al. (2024) introduced the KM framework AGENTiGraph for natural language interaction to enhance KG-user interaction efficiency. The framework integrates knowledge extraction, integration, and real-time visualization to improve user experience, and comprises seven agents that convert natural language into structured graph operations. Additionally, it provides a pre-designed functional framework to support diverse user needs in KG interactions and exploration. Experimental evaluations showed the framework achieved 95.12% accuracy in question classification and 90.45% in question execution, with user surveys demonstrating strong applicability in cross-domain KM tasks. However, the study did not explicitly address how data type diversity (combining structured and unstructured data) impacts system performance, nor did it validate the framework in high-context-dependent fields like accounting or supply chain management.

## 2.2. Leveraging Knowledge Graphs and Large Language Models for Knowledge Management in SMEs

A growing body of research is dedicated to developing efficient and cost-effective KM solutions, particularly through the integration of KGs and LLM. This approach addresses the challenge faced by SMEs in adopting high-cost technologies due to limited resources.

### 2.2.1. Application of KG for Supply Chain KM

Mo et al. (2023) introduced an innovative and cost-efficient KM solution that integrates knowledge graphs (KGs) with AI techniques, combining Neo4j with

LLM-powered visual systems to support KM activities. This groundbreaking solution leverages cost-effective LLM visual processing technology and KG integration to achieve powerful KM and recommendation capabilities, aiming to revolutionize data interpretation and decision-making processes in SMEs' supply chains, particularly in manufacturing, enabling real-time and informed operational choices. SMEs can ensure cost savings by leveraging the system's flexibility and KG insights. However, the study primarily focuses on visual recognition data and has not yet delved into text knowledge extraction and semantic reasoning capabilities.

Mansfield et al. (2021) utilized UML, a widely adopted ontology modeling tool for constructing KG and facilitating effective communication, which is sufficiently abstract to be comprehensible to users. This study addresses challenges faced by SMEs in adopting semantic technologies to organize and leverage their data, including issues such as insufficient human and financial resource allocation, lack of comprehensive data governance policies, and organizational resistance to new practices and knowledge. However, the article does not discuss practical methods for creating knowledge graphs using LLM, and the semi-automated ontology construction method remains time-consuming despite achieving high accuracy levels.

Galkin et al. (2017) introduced the concept of Enterprise Knowledge Graph (EKG) and defined its position within ERP architecture. As a carrier integrating corporate supply chain and financial information, ERP systems also achieve the function of connecting supply chain and financial operations (Lukyanova, Haddud, & Khare, 2022, Halimuzzaman & Sharma, 2023). EKG occupies an intermediary position between the original enterprise data storage layer and the ontology mapping layer. The ontology mapping layer contains a set of ontologies that provide mapping relationships between knowledge graphs and ontologies. The study proposes an independent evaluation framework for comparing current KM solutions with EKG integration. Observations indicate that current state-of-the-art methods fail to fully support EKG, meaning no KM solution can simultaneously meet all the evaluation dimensions of EKG functions. This can be understood as the reality of extending traditional data architectures to semantic data architectures as fully customized enterprise tools being impractical in practice. While this study explores KM based on KG, successfully marring KG with ontologies conceptually and defining role of KG in ERP architecture, the KM solutions presented employ diverse KG creation techniques. Although one solution incorporates NLP, it lacks in-depth analysis of cutting-edge technologies and their practical effects, such as addressing potential ambiguities in KG. Consequently, the research primarily focuses on examining user experience, data quality, and the conceptual implementation of this integration within existing KM solutions in ERP architecture.

#### 2.2.2. Application of KG for Accounting KM

Chen & Wang (2025) explored an automated accounting KG construction method integrating GPT-4 and prompt engineering. The results demonstrated that this model successfully extracts rich knowledge pentuples from large volumes of unstructured accounting data, enabling the construction of accounting knowledge graphs that lay the foundation for querying and visualizing knowledge. This study proves the feasibility of BERT-based accounting KG construction, significantly improving efficiency. However, the generalization capability of the model for non-standard documents (e.g., work notes) in real-world accounting practice requires further validation, since experimental data primarily consisted of standardized legal texts. Additionally, while the pentuple knowledge structure is more detailed and clear, it may require greater storage space, posing challenges for SMEs with limited IT resources.

To support knowledge application in accounting tasks such as bookkeeping, Schulze et al. (2022) proposed a process for constructing and enriching KG for accounting from heterogeneous accounting resources. The feasibility of this approach was demonstrated through practical application using real corporate data. A series of prototype knowledge services were implemented, including recommending similar accounting cases to accountants by harnessing the accounting knowledge graph. The study revealed that using semantically rich data from the knowledge graph yields better results compared to non-semantically enriched data. A key highlight of this research is the diversified data sources in the KG, enabling experimental scenarios to closely resemble real-world accounting practices. However, the study did not employ LLM-based architectures for KG formation, limiting its reference value.

In conclusion, existing research highlights the strong potential of knowledge graphs and large language models in KM. Nevertheless, three gaps remain in the context of small and medium-sized enterprises (SMEs):

1) Lack of portability and cost-benefit verification for low-resource enterprises;

2) Multi-focus vision or structured data, lack of unstructured text processing mechanism;

3) Current research primarily focuses on sectors like healthcare and law, while empirical studies on the integration of accounting and supply chain applications remain limited.

## 3. Research Method

### 3.1. Research Design

Rather than developing new deep learning architectures, this study employs a mature BERT-based LLM for unstructured information extraction and uses Neo4j to construct the KG. This integrated approach enables the study to meet its research objectives and substantially improves decision-making efficiency in the largely underexamined areas of SME accounting and supply chain management.

The study generally follows three phases:

1) Data preprocessing: A medical equipment manufacturer is selected as the SME case study, whose accounting and supply chain business documents is collected and used the text data as the input dataset, followed by data annotation and merge conversion. The output dataset serves as the training data for the models,

enabling automatic extraction of entities and relationships from the text in subsequent stages;

2) Model training: First, training parameters are configured, while the training dataset is delivered from the previous stage to the model training tasks, namely the Entity Recognition (NER) architecture and the Relation Recognition (RE), in order to implement text entity and relation judgment and prediction. Each model includes two different combinations of models using BERT as the embedding layer, aiming to identify more accurate models based on the automatic extraction indices of entities and relations calculated during model training, to be used for the next phase of KG formation;

3) KG creation: Entities and relationships are identified, extracted and predicted, predict relationships by integrating the real data set and using a relatively accurate NER model and a RE model, so as to build triple entity-relation-entity, and finally connect with Neo4j to store the knowledge graph for visualization and query.

The proposed framework utilizes deep learning algorithms to identify, predict, and extract textual entities and their relationships through exploratory and developmental innovation in KM, hence constructing a KG-based repository. The accuracy of each model is evaluated through BERT series model training, with parameters calculated by comparing the results against manually annotated data.

## 3.2. Baseline Models

This research harnessed end-to-end method, by which NER and RE tasks are implemented independently.

### 3.2.1. NER Model

This study considers two supervised methods as candidate NER models, which will be evaluated in terms of prediction accuracy. Specifically, the model training is implemented using the PyTorch framework, with evaluation metrics being generated and the optimal model being saved. Then the outperformer will be selected as the baseline method for NER. This is discussed in Section 4.2. These candidate methods are outlined below.

**BERT + CRF**

In this composite model architecture (see **Figure 1** and **Appendix**), the BERT layer serves as a feature extractor responsible for capturing rich contextual semantic features, while the CRF layer acts as a decoding layer to optimize label transition probabilities and enhance entity recognition accuracy. Currently, models like BERT-CRF and BERT-Bi-LSTM-CRF are widely adopted as baselines for many named entity recognition tasks, achieving excellent performance. This is largely attributed to BERT's powerful contextual representation capabilities, simplified feature engineering, cross-domain generalization ability, and rapid model adaptation, which have been extensively applied in NLP tasks (Devlin et al., 2019). Additionally, the NER model employs a sequence annotation approach supporting multiple annotation schemes (e.g., BIO, BIOES), enabling integration with BERT-based machine learning frameworks (Zhu et al., 2020; Park et al., 2022).

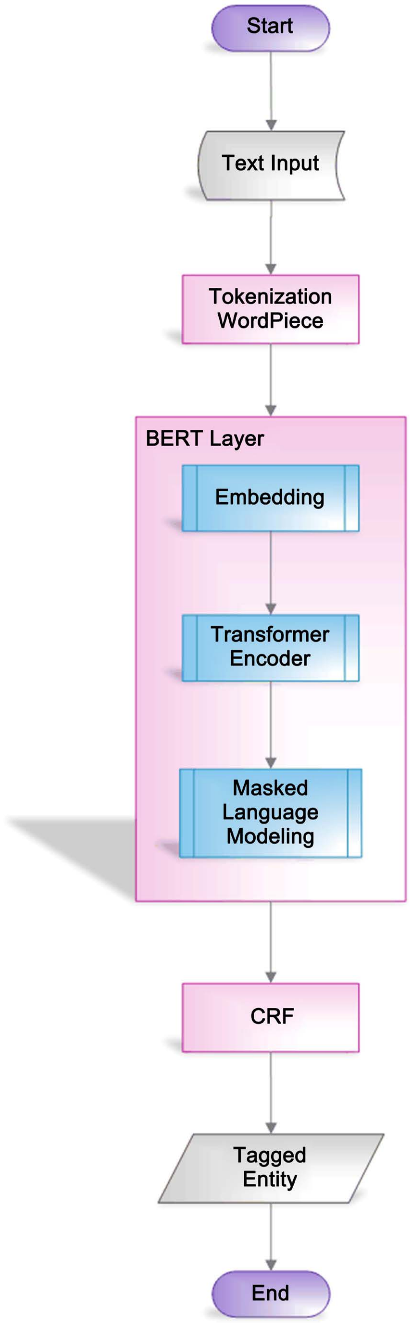


**Figure 1.** BERT + CRF model.

The output from the BERT layer is then fed into the Conditional Random Field (CRF) layer. Sequence-based named entity recognition methods typically employ models like BERT to encode text token sequences, followed by final label determination through CRF. The CRF layer optimizes annotation results by considering word relationships within sentences and transition probabilities between labels (Sutton & McCallum, 2012). This allows the CRF layer to correct errors in initial word classification by leveraging adjacent word labels. In this study's case, the CRF layer recognizes that "Debit" typically appears between accounting items and monetary amounts, thereby increasing the probability of correct labeling. CRF ensures generated label sequences adhere to the B-I-O tagging rules, preventing invalid sequences. Consequently, scaling the CRF learning rate by a factor of 100 significantly improves model performance, demonstrating that CRF requires a higher learning rate than BERT, and that increasing the learning rate enhances model performance.

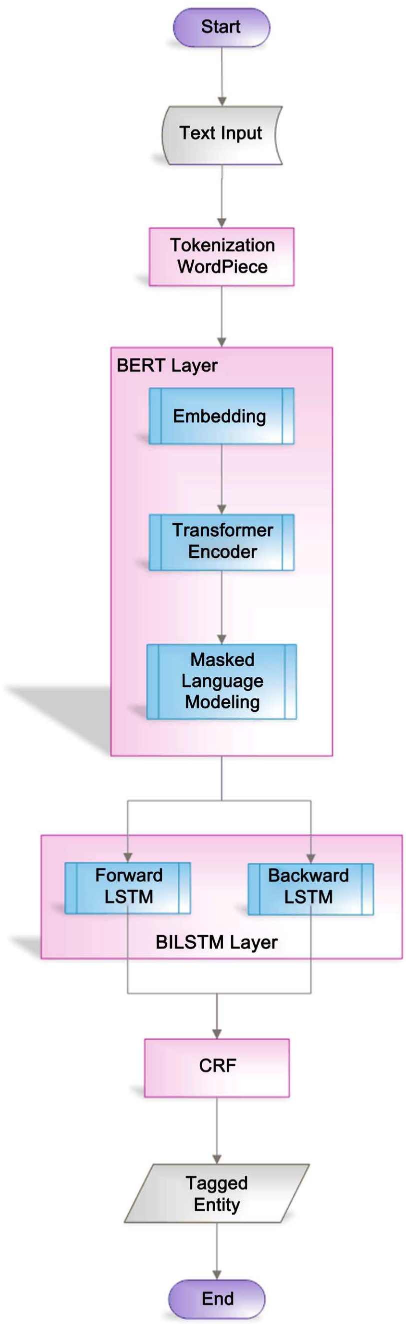


Figure 2. BERT + BiLSTM + CRF model.

**BERT + BiLSTM + CRF**

In this composite model architecture (see **Figure 2** and **Appendix**), both the BERT and CRF layers retain the same functionalities as the previous BERT + CRF model. Its unique advantage lies in integrating BiLSTM capabilities, enabling it to learn features of each word in sentences from two directions. It effectively maintains information flow across long sequences with specialized memory cells, ultimately accurately identifying named entities based on contextual information. This bidirectional contextual understanding combined with robust label sequence modeling allows BILSTM + CRF to outperform standalone LSTM or CRF models in extracting and classifying technical terms in texts (Huang, Xu, & Yu, 2015). In the case of the study, the BI-LSTM-CRF model can simultaneously consider contextual information such as the business activity "Purchase Invoice CGRK20230718003" and the accounting subject name "Raw Materials" when processing the accounting code "10302". After this step, the model generates preliminary classification results for each word—for example, "10302" might be identified as ACCOUNTCODE, "Purchase Invoice CGRK20230718003" as BUSINESSACTIVITY, and "Raw Materials" as ACCOUNT. Meanwhile, the CRF layer ensures consistent entity tagging across the entire sequence, preventing account names from being mislabeled as product names.

#### 3.2.2. RE Model

This study considers two supervised methods as candidate RE models, which will be evaluated in terms of prediction accuracy. Specifically, the model training is implemented using the PyTorch framework, with evaluation metrics being generated and the optimal model being saved. Then the outperformer will be selected as the baseline method for RE. This is discussed in Section 4.2. These candidate methods are outlined below.

**BERT**

In entity relation extraction tasks, the input typically consists of a sentence requiring judgment and two entities. A common embedding method involves calculating positional vectors of the two entities within the sentence to annotate them. First, a multi-label classification model is used to determine the types of relationships in the sentence, identifying how many relationships exist in a single statement. Second, the sentence and its predicted relationship types are fed into a sequence labeling model. For example, if the sentence predicts three relationships, it is split into three samples. Next, entity recognition is performed to identify entities based on the sequence labels. Finally, combining the predicted relationships and entities, a triple (entity1, relation, entity2) is generated. Taking the example in this study: an input sentence like "ERP system notification will be sent to the warehouse staff" and a pair of entities identified through NER tasks, such as "ERP system notification" and "warehouse staff". The output is the semantic relation triple between entities ("warehouse staff", "recipient_of", "ERP system notification").

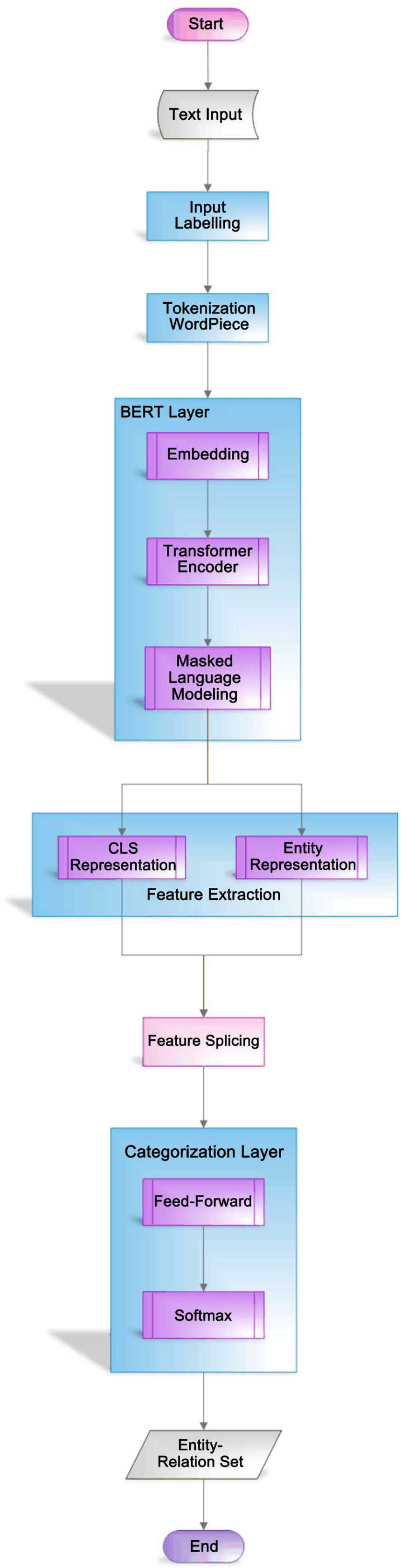


**Figure 3.** BERT model.

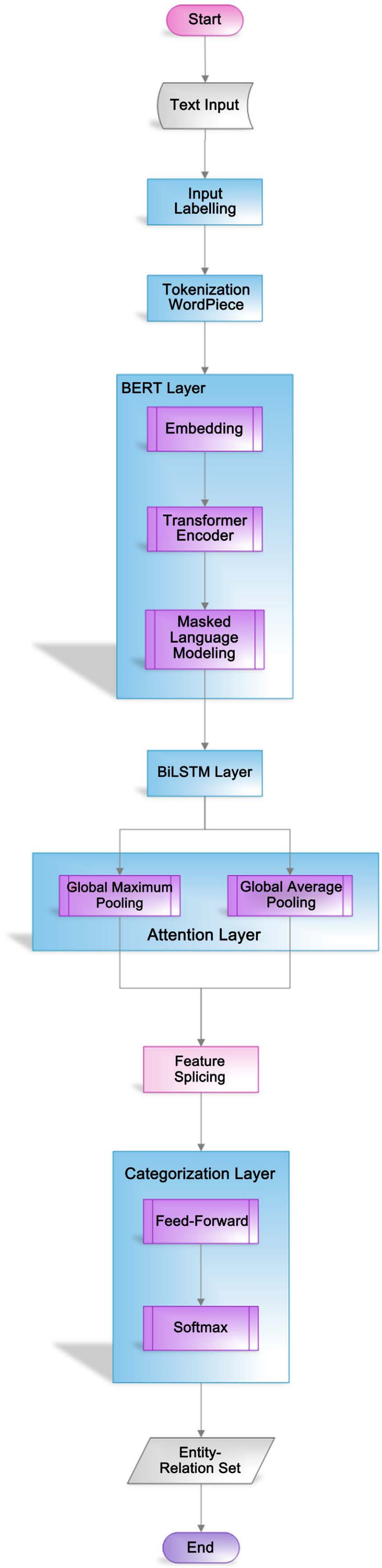


**Figure 4.** BERT-BiLSTM model.

In this study's RE model (see **Figure 3** and **Appendix**), special tags are applied to entities in the input layer, providing boundary and positional information for each entity pair. Meanwhile, the BERT layer not only adds relative position encoding to the text to enhance dependency understanding but also utilizes entity pair embedding and [CLS] sentence separators. This approach captures both sentence-level features and semantic features of entities, significantly improving model performance (Giorgi et al., 2019; Wu & He, 2019). Unlike NER models, CRF is not introduced here, as relation extraction is typically considered a sentence-level classification task.

**BERT + BiLSTM**

The model is the hybrid of BERT, BILSTM, and attention mechanisms (see **Figure 4** and **Appendix**). BiLSTM + Attention stands as a cornerstone structure in relation classification tasks, renowned for its robust contextual modeling capabilities and strong interpretability. It serves as a crucial foundation for developing sophisticated deep relational extraction systems. Here, BERT delivers powerful text representation through context-aware word vectors, BiLSTM extracts contextual information, while Attention mechanisms prioritize key entities by assigning weighted importance to relevant terms (Zhou et al., 2016; Wei, Xu, & Hu, 2021). This study innovatively optimizes the attention layer into a pooling attention mechanism. With the max-pooling or average-pooling operations, the model aggregates attention weights to enhance feature recognition capabilities.

#### 3.2.3. Model Configuration

This study defines a series of hyper-parameters to configure the training process of an NLP model based on BERT + BiLSTM + CRF (**Table 1**). Thereby, the experiment adopts a typical two-stage optimization setup used in the NLP tasks. The parameters that are set for both NER and RE tasks are outlined in the **Table 1**. It is noted that the BERT component uses a relatively small learning rate (3e−5)

**Table 1.** Model configuration parameters.

| Task / Parameter | NER | RE |
|---|---|---|
| Max_seq_len | 256 | 256 |
| Epochs | 50 | 50 |
| Train_batch_size | 24 | 24 |
| Dev_batch_size | 12 | 12 |
| Bert_learning_rate | 3e−5 | 3e−5 |
| CRF_learning_rate | 3e−3 | - |
| Adam_epsilon | 1e−8 | 1e−8 |
| Weight_decay | 0.01 | 0.01 |
| Warmup_proportion | 0.01 | 0.01 |
| Save_step | 20 | 20 |

to fine-tune the pre-trained model, while the CRF layer employs a large learning rate (3e−3) to rapidly converge the sequence labeling head. An Adam optimizer, L2 regularization strategy and linear warming mechanism are also integrated into the model so as to improve the stability. To mitigate the risk of overfitting, a combination of the weight decay for restricting weight norm, the warm-up mechanism ensuring smooth start and the appropriate batch sizes forming a complete regularization system is deployed. Notably, the design of decreasing batch sizes (train > dev) is commonly used for stable training in resource-constrained scenarios.

## 4. Research Process & Implementation

### 4.1. Data Pre-Processing

**Source**: Take a small and medium-sized enterprise in medical device manufacturing as a case, collect its accounting and supply chain related texts (such as voucher/document description, ERP operation instructions, purchase/sales records, etc.).

**Data cleaning**: Remove elements that are irrelevant to the field or involve privacy (names, addresses, etc.), and categorize them by business domain (such as "supply chain-sales", "finance-vouchers", "ERP-purchasing").

**Annotation Tools and Formats**: After data cleaning, we utilized Doccano—a machine learning text annotation tool—to generate two types of labels: entity tags and relation tags. Each text was manually annotated by adding tags to describe entities and relationships, forming triplets of entities and relations for the LLM to generate target responses. Figure 5 and Figure 6 demonstrate an example of text content annotation, triplet extraction, and the corresponding JSONL format scripts for entities and relations. Finally, the Python command "merge_jsonl_files()" was used to merge all annotated files into a unified dataset.

The processed data set is sectioned into two subsets based on the ratio of 8:2, that is, 80% of the data (training subset) is exploited for model training, and the rest is the held-out set aiming for model accuracy evaluation.

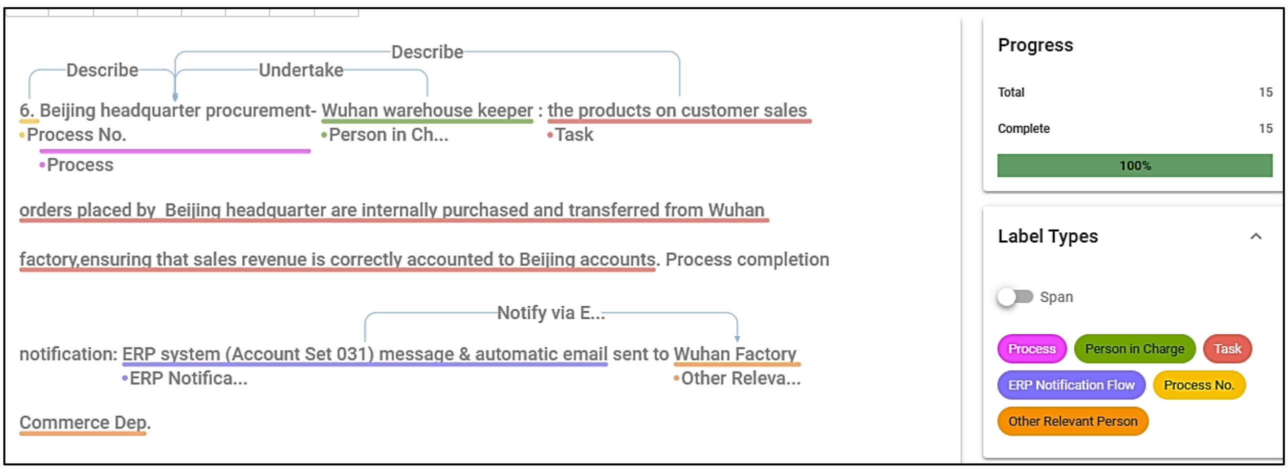


Figure 5. Text content annotation.

```
15 {"id":268,"text":"6. Beijing headquarter procurement- Wuhan warehouse keeper : the products on customer sales orders placed by  Beijing headquarter are internally purchased and transferred from Wuhan factory,ensuring that sales revenue is correctly accounted to Beijing accounts. Process completion notification: ERP system (Account Set 031) message & automatic email sent to Wuhan Factory Commerce Dep.","entities":[{"id":723,"label":"Process","start_offset":3,"end_offset":34},{"id":724,"label":"Process No.","start_offset":0,"end_offset":3},{"id":725,"label":"Person in Charge","start_offset":36,"end_offset":58},{"id":726,"label":"Task","start_offset":61,"end_offset":260},{"id":727,"label":"ERP Notification Flow","start_offset":295,"end_offset":349},{"id":728,"label":"Other Relevant Person","start_offset":358,"end_offset":384}],"relations":[{"id":568,"from_id":727,"to_id":728,"type":"Notify via ERP"},{"id":569,"from_id":724,"to_id":723,"type":"Describe"},{"id":570,"from_id":726,"to_id":723,"type":"Describe"},{"id":571,"from_id":725,"to_id":723,"type":"Undertake"}],"Comments":[]}
```


Figure 6. JSONL format scripts for entities and relations.

### 4.2. Model Training and Evaluation

An end-to-end where both NER and RE each contained two LLM-BERT-based models is employed in model training phase, which aims to evaluate the performance and accuracy of each model, identifying the most effective one for the subsequent KG creation. Metrics such as Precision, Recall, and F1-score are used to assess the performance of the four baseline models. Higher values indicate better model performance. This section presents the model performance on the cleaned dataset compared to the baseline models, with all metrics displayed as weighted average scores calculated based on sample size for each category, as outlined in Table 2.

Table 2. Model evaluation performance.

| Metrics / Model | Precision | Recall | F1-score |
|---|---|---|---|
| NER − BERT + CRF | 0.66 | 0.74 | 0.69 |
| NER − BERT + BiLSTM + CRF | 0.73 | 0.76 | 0.74 |
| RE − BERT | 0.69 | 0.80 | 0.74 |
| RE − BERT + BiLSTM | 0.65 | 0.76 | 0.70 |

Just as shown in Table 2, all the BERT and its portfolio models in the study render good performance regarding NER and RE tasks, with all metrics figures stand above 50%. Comparing the NER model evaluation results, the BERT-based hybrid model BERT + BiLSTM + CRF outperformed most sample prediction tasks, resulted in evidently higher rates for all of the three metrics Precision (0.73), Recall (0.76) and F1 (0.74) than the other NER model with Precision (0.66), Recall (0.74) and F1 (0.69) (see Table 2 & Figure 7). Regarding the RE models, the sole BERT model did better on the held-out set than the model BERT + BiLSTM (see Table 2 & Figure 8). As the former obtained relatively greater metrics for Precision (0.69), Recall (0.80) and F1 (0.74) than those of the latter.

In model selection, the system automatically evaluates performance on the held-out set during training and retains the model with the highest F1 score. Based on the SMEs in the knowledge economy context, the final model selection must balance accuracy and completeness—ensuring minimal omission of entities and relationships—where all three metrics carry equal weight. According to the analysis above, the NER model BERT + BiLSTM + CRF and the RE model BERT are chosen to perform the construction of KG.

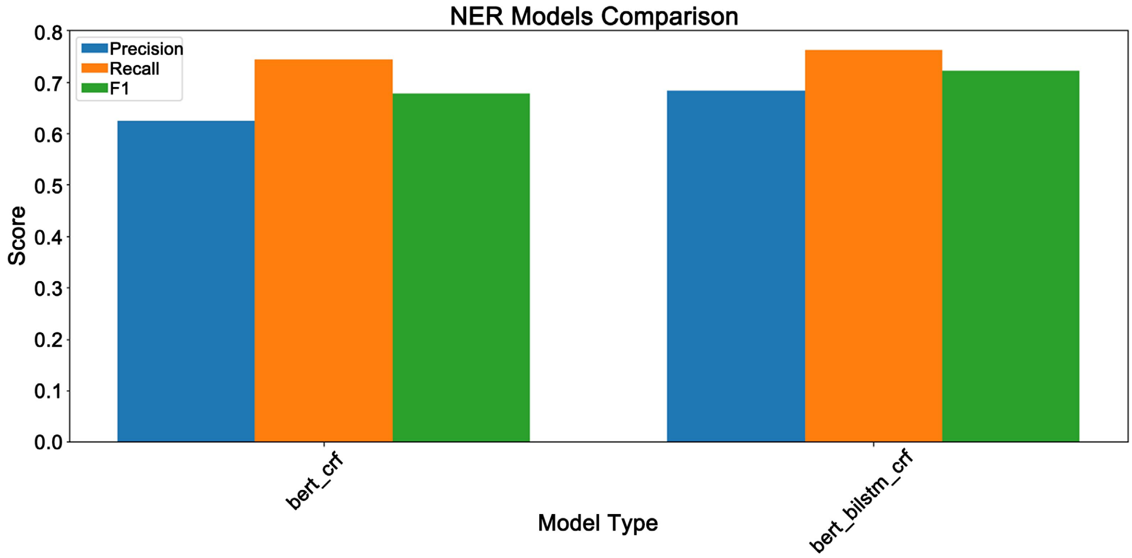


**Figure 7.** NER model evaluation comparison.

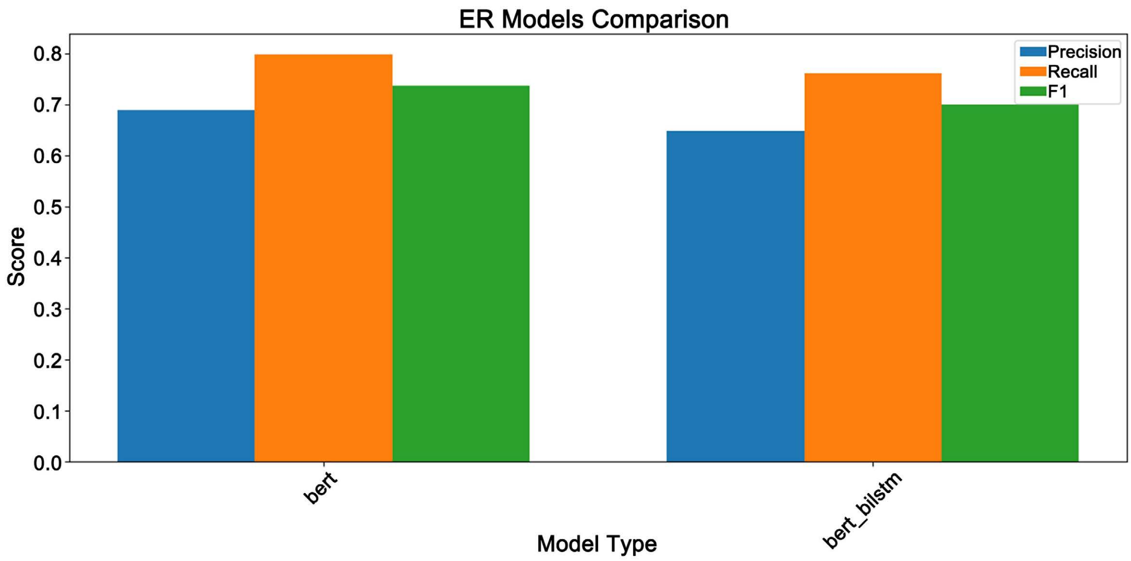


**Figure 8.** RE model evaluation comparison.

## 4.3. Knowledge Graph Construction

In this experiment, the industry-leading cloud database service Neo4j Aura is harnessed to visualize and manage the KG. The collected data will be imported into the Neo4j graph database, which can effectively improve data readability and analytical efficiency. The graph not only comprehensively presents the knowledge in accounting and supply chain of SMEs studied in this paper but also helps us identify patterns in the data, providing key insights for optimizing decision-making.

In addition, Py2neo tool is applied to attempt seamless integration between Neo4j, databases, and software. This tool supports Python-based automation for updating knowledge graphs in Neo4j, perfectly aligning with our proposed automated configuration solution. The KM system is dedicated to optimizing leakage detection processes across various products. The core component is a BERT-based RE and NER predictor that seamlessly interfaces with the cloud-based Neo4j graph database. The formation of the KG primarily involves three phases, particularly suitable for automated end-to-end workflows requiring joint cooperation of NER and RE tasks, as detailed below:

1) Configuration—The pre-training models are loaded and the environment parameters are configured simultaneously, in order to process the input text to meet the requirements of the model, and finally carry out inference and prediction;

2) Prediction—NER and RE predictors are utilized to predict entities and relationships in the input text, followed by the entities extraction processes—truncate sentences → add special markers → convert to ID sequences → pad to fixed length. The processed data is fed into the preloaded BERT + BiLSTM + CRF structure NER model for decoding for prediction of pairwise relationships, requiring additional input of head and tail entity information to assemble into an input format suitable for the RE model. These are finally encapsulated into a unified interface for external calls;

3) Construction—The prediction resulted from the previous phase are imported into the Neo4j graph database to build a KG by employing Py2neo. The process involves parsing entities and relationships from the imported data, creating corresponding entity nodes and relationships, and recording the import count.

### 4.4. Result & Analysis

The results of this study show that the structured text information in JSON format, namely entities and relationships, which are inferred and recognized by LLM, are converted into nodes and connecting lines in Neo4j graphs by using Py2neo. The links and constructions between them generate RDF-like triple-form KG, laying the foundation for knowledge querying and visualization. Specifically, each node represents subjects or objects, with each connecting line being predicates. The graph also has the following two characteristics, further enhancing the accuracy of it:

- Node uniqueness guarantee: The merge() method checks for duplicate nodes using the composite primary key (label + name), ensuring no duplicate creation;
- Directionality: Clearly distinguish between source nodes (from_id) and target nodes (to_id) to align with directed graph semantics.

After completing the schema construction, Neo4j's built-in visualization tools enable direct database connection for dynamic data presentation. Specific entities and their associated relationships can be rapidly retrieved through Cypher queries in Neo4j. The platform supports node relationship queries. Taking the "Purchase Receipt" entity as an example, the Cypher query results (**Figure 9**) demonstrate multi-entity relationships, revealing links among accounting entities and a supply chain entities. The yellow nodes represent accounting entities with the remaining nodes being those of supply chain. This visual representation clarifies complex networks, facilitating quick identification of interconnections between SMEs' accounting entities and key supply chain components.

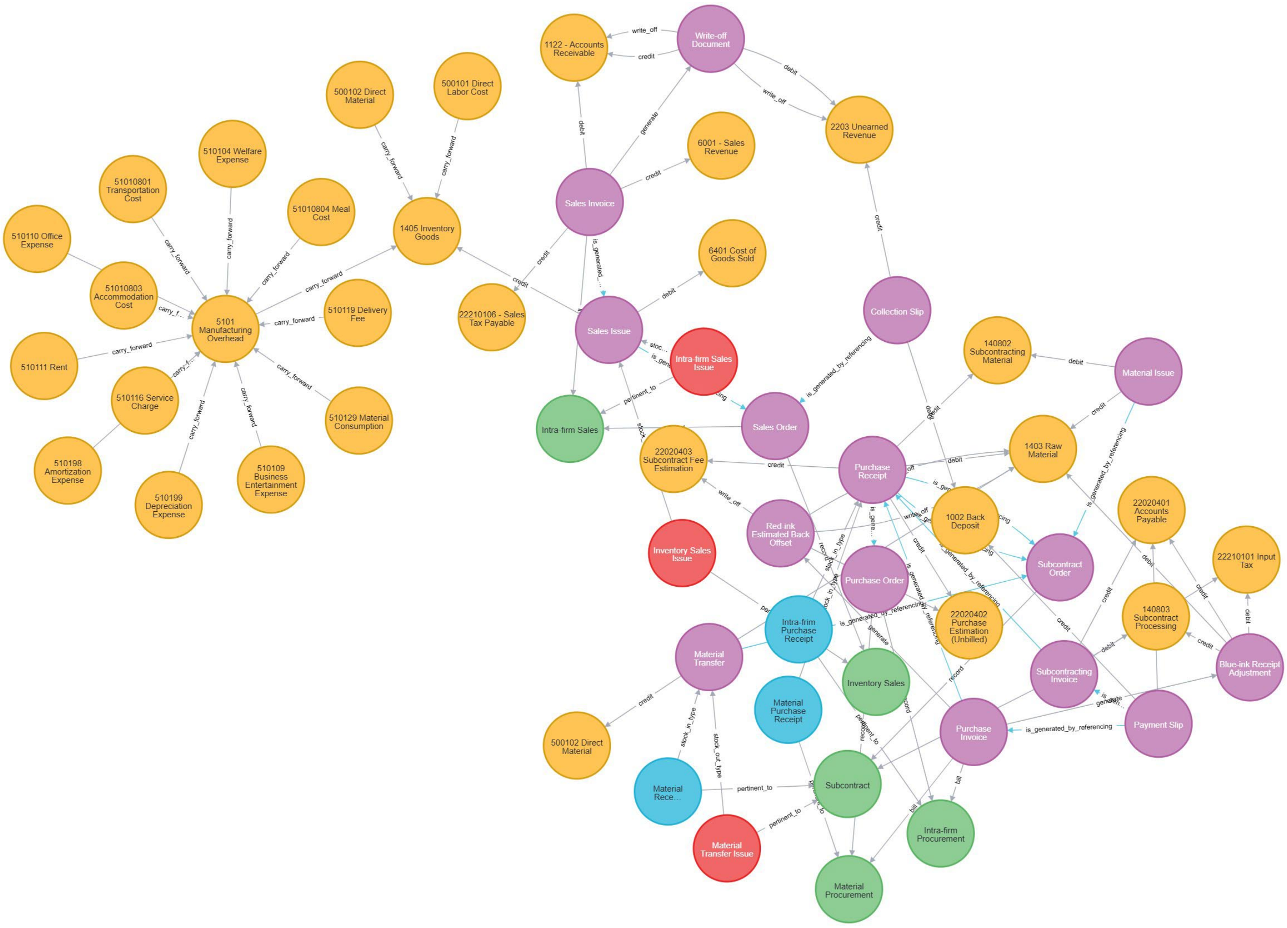


**Figure 9.** Knowledge graph visual query on Neo4j.

## 5. Discussion

The training and evaluation results demonstrate that the models achieved competitive performance in both NER and RE tasks. The precision levels of the models employed in KG construction reached approximately 70%, with F1-scores as high as 74% and recall rates reaching 80%. This is attributed to BERT for the effective handling of multi-structure contexts, and BiLSTM for the enhanced text learning capabilities, precise input labeling, and the deeper semantic understanding enabled by CRF and attention mechanisms, all of which collectively improve prediction accuracy. The evaluation results demonstrate that LLM-based KG creation can effectively and accurately identify entities and relationships. With the robust capabilities of the LLM models, this approach not only lays the foundation for accurate accounting and supply chain KG systems, but also has potential for cost-effectiveness (e.g. reduced manual effort) in KM realization for SMEs because of the automation of KG formation harnessing cost-saving and open-source implementation tool PyTorch that makes creating and training the models flexible and simple. Furthermore, the interaction between LLM and KG is not one-way but mutually reinforcing. Current research indicates that LLMs exhibit limitations in accounting, particularly in reasoning capabilities. KG can compensate for these deficiencies, thereby establishing a foundation for developing specialized models in the accounting domain (Zhong, Chen, & Wang, 2024).

The knowledge on accounting and supply chain for SMEs is effectively managed, facilitating knowledge organization, discovery, and sharing. In many cases, SMEs faces the issues of data quality and inconsistency encountered with the ERP and accounting and supply chain texts, such as fragmented data sources, terms and concepts that needs explanation for new users and lack of knowledge transfer (e.g. mentoring). By regularly collecting and updating data from multiple authoritative internal and external sources, the system automatically stores data into a KG to ensure accuracy and timeliness. The search function of KG enables accounting and supply chain personnel to quickly identify relevant information. Leveraging relational information in the financial and tax knowledge graph, it promotes connections between different financial, tax, and supply chain knowledge, uncovering potential correlations and providing a more comprehensive knowledge perspective. Knowledge sharing and collaboration are achieved through the KG capability, enhancing knowledge exchange among team members and significantly improving the efficiency and effectiveness of KM.

Moreover, KG significantly enhances the accuracy and timeliness of financial and supply chain information queries. This study evaluates and tests questions regarding information inquiries, yielding precision metrics. The findings can be applied in KG-based KM systems, enabling users to receive personalized financial management and compliance recommendations while boosting automation in corporate internal audits and risk control. An interactive KM system calculates candidate answers based on content similarity. The most matching answer is then selected and delivered to financial and supply chain personnel for final resolution. This provides intelligent search and Q&A capabilities, ensuring staff can swiftly access accurate relevant knowledge and improve efficiency in accounting and supply chain information queries.

Also, systematic and personalized training for accounting and supply chain professionals with the assistance of KG, enhancing learners' motivation and practical application skills while improving knowledge transfer efficiency. The SME-focused financial and supply chain KG constructs tax and accounting frameworks covering regulations, accounting principles, and supply chain operations. It visually represents entities (business units, accounting items, documents), attributes (definitions, scopes, update dates), and relationships (dependencies, inclusion rules), making complex knowledge instantly understandable. Customized training pathways are developed based on roles, experience, and educational backgrounds of employees. The relationships of KG enable targeted recommendations of specialized knowledge points, ensuring precise training delivery. Real-time updates maintain the relevance of the graph, guaranteeing employees access to the latest tax and financial information. A notification system promptly alerts staff about critical updates. Finally, interactive learning tools are created using the KG. Employees can explore nodes to gain in-depth understanding, while interactive Q& A sessions ensure effective knowledge transfer.

In addition, graph databases like Neo4j are gaining increasing industry traction due to their exceptional capabilities in processing complex data and simplifying data analysis and visualization. The platform's reliability and secure architecture ensure both accessible and protected data. As another key advantage of cloud services, Neo4j Aura offers cost-effectiveness. The cloud service model requires no substantial infrastructure investment and little subscription costs, making it affordable even for SMEs with limited budgets, offering cutting-edge database solutions within reach and allowing SMEs to adopt advanced technologies without heavy capital expenditure. More importantly, it enables high-quality outcomes at reduced costs. SMEs can effectively ensure product quality and customer satisfaction—crucial for success in highly competitive markets by utilization of system flexibility and intelligent KG analytics (Mo et al., 2023).

## 6. Conclusion

This study employs BERT-based LLM to conduct in-depth research on KG and KM for SMEs in the accounting and supply chain sectors, while automating the acquisition, organization, and updating of relevant data for SMEs' accounting and supply chain KM, thereby constructing KG applying implementing LLM-based automation. First, unstructured data undergo intelligent processing. Subsequently, hybrid models based on BERT for NER and RE are developed, tested, and evaluated. These models are deployed to construct KG, intelligently capturing and presenting knowledge with an F1-score exceeding 70%. The NER and RE models automatically constructs "entity-relation-entity" RDF triple structures from processed texts, mapping entities and with their interrelations. The knowledge derived from LLM analysis is then stored and visualized as a KG using Neo4j, a high-performance and cost-effective graph database. This KG construction approach provides both technical and practical foundations for developing an LLM-enabled KG-based KM system.

The findings of this study effectively empower users to access knowledge in an intelligent, efficient, professional, and streamlined manner, significantly enhancing work efficiency and quality. This research provides a possible scholar support for a knowledge-learning solution that enhances the "intelligence" of KM through extensive human-assisted and machine-driven Q& A learning. The semantic search capabilities of the KG enable an interactive system integrated with cost-effective yet high-performance graph databases like Neo4j. This will optimize information retrieval efficiency, providing employees with knowledge aggregation, real-time updates, precise searches, and AI-powered Q& A support. Additionally, the KG will be utilized to develop a training platform offering intelligent training for new employees, accelerating knowledge transfer and learning. In summary, SMEs can gain deep insights into their operational processes, make informed decisions, and thrive in the rapidly evolving digital economy via harnessing the powerful capabilities of LLM and KG technology.

KG holds significant value for promotion as KG-enhanced models improve understanding of accounting terminology, supply chain policies, and industry-specific knowledge, enabling more accurate support for tasks such as information extraction, document review, and risk identification, thereby achieving more precise retrieval capabilities. This study provides a data foundation for optimizing KM frameworks for SMEs and paves a new path for efficient construction and mining in the KG domain.

## 7. Limitation

The limitations are as follows:

- Resource consumption: BERT model pre-training requires a large amount of computing resources and high hardware requirements;
- Data dependency: BERT model requires a sufficiently large annotated dataset for fine-tuning to achieve optimal performance. Moreover, the construction of KG depends on the quality and comprehensiveness of the initial data. If the data is biased or insufficient, it may affect the accuracy and reliability of the graph;
- Black box characteristics: The pre-training process of BERT model is relatively opaque, and the model is less interpretable than the traditional model based on rules;
- RE model: It cannot extract multiple triples simultaneously (unable to handle triple overlap issues), has fixed input length (unable to adapt to long text relationships for cross-sentence reasoning), is only suitable for sentence-level classification, and cannot perform end-to-end joint extraction;
- Industry terms and complex entity relationships: LLMs may not fully capture the inherent meaning of specific industry terms and complex financial relationships, although they have good performance in extracting information;
- KG maintenance: Maintaining the dynamic updates of KG presents a challenge, which requires onging refinement of data sources and update mechanisms to ensure the graph's timeliness;
- Generalizability: The scope of the study is limited on a single case SME, thus future work are needed to validate the LLM-based KM framework across different SME sectors.

## Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

## References

Al Koliby, I. S., Noor, N. H. M., Al-Swidi, A. K., Al-Hakimi, M. A., & Mehat, N. A. B.

(2025). Enhancing Sustainable Performance among Manufacturing SMEs: The Interplay of Knowledge Management and Organizational Structure. *Discover Sustainability, 6,* 1-16. https://doi.org/10.1007/s43621-025-01351-1

Chen, W., & Sun, T. (2024). Application of Artificial Intelligence with Knowledge Graph Technology in Enterprise Financial and Tax Knowledge Management. *Financial Supervision, 5,* 101-104.

Chen, S. S., & Wang, M. (2025). Construction and Application Prospect of Accounting Knowledge Graph Based on Large Language Model. *Friends of Accounting, 5,* 152-161.

Dehshiri, S. J. H., Amiri, M., Mostafaeipour, A., Pamučar, D., & Le, T. (2024). Enhancing Supply Chain Performance by Integrating Knowledge Management and Lean, Agile, Resilient, and Green Paradigms. *Journal of Management Analytics, 11,* 738-769. https://doi.org/10.1080/23270012.2024.2408527

Devlin, J., Chang, M., Lee, K., & Toutanova, K. (2019). *BERT: Pre-Training of Deep Bidirectional Transformers for Language Understanding.* North American Chapter of the Association for Computational Linguistics.

Du Plessis, T., & du Toit, A. S. A. (2006). Knowledge Management and Legal Practice. *International Journal of Information Management, 26,* 360-371. https://doi.org/10.1016/j.ijinfomgt.2006.06.003

Durst, S., Edvardsson, I. R., & Foli, S. (2023). Knowledge Management in SMEs: A Follow-Up Literature Review. *Journal of Knowledge Management, 27,* 25-58. https://doi.org/10.1108/jkm-04-2022-0325

Eslami, M. H., Achtenhagen, L., Bertsch, C. T., & Lehmann, A. (2023). Knowledge-Sharing across Supply Chain Actors in Adopting Industry 4.0 Technologies: An Exploratory Case Study within the Automotive Industry. *Technological Forecasting and Social Change, 186,* Article ID: 122118. https://doi.org/10.1016/j.techfore.2022.122118

Galkin, M., Auer, S., Vidal, M.-E., & Scerri, S. (2017). Enterprise Knowledge Graphs: A Semantic Approach for Knowledge Management in the Next Generation of Enterprise Information Systems. In *Proceedings of the 19th International Conference on Enterprise Information Systems* (Volume 2, pp. 88-98). SCITEPRESS-Science and Technology Publications. https://doi.org/10.5220/0006325200880098

Giorgi, J., Wang, X. D., Sahar, N., & Shin, W. Y. (2019). *End-to-End Named Entity Recognition and Relation Extraction Using Pre-Trained Language Models.* 1-12.

Halimuzzaman, M., & Sharma, J. (2023). The Evolution of Accounting Information Systems (AIS) and Enterprise Resource Planning (ERP): A Review of Literature. *Episteme: An Online Interdisciplinary, Multidisciplinary & Multi-Cultural Journal, 12,* 170-190.

Hu, X., Liu, A., & Dai, Y. (2024). Combining ChatGPT and Knowledge Graph for Explainable Machine Learning-Driven Design: A Case Study. *Journal of Engineering Design, 36,* 1479-1501. https://doi.org/10.1080/09544828.2024.2355758

Huang, Z.H., Xu, W., & Yu, K. (2015). *Bidirectional LSTM-CRF Models for Sequence Tagging.*

Li, J., Qian, L., Liu, P., & Liu, T. (2024). Construction of Legal Knowledge Graph Based on Knowledge-Enhanced Large Language Models. *Information, 15,* Article No. 666. https://doi.org/10.3390/info15110666

Lukyanova, I., Haddud, A., & Khare, A. (2022). Types of ERP Systems and Their Impacts on the Supply Chains in the Humanitarian and Private Sectors. *Sustainability, 14,* Article No. 13054. https://doi.org/10.3390/su142013054

Mansfield, M., Tamma, V., Goddard, P., & Coenen, F. (2021). Capturing Expert Knowledge for Building Enterprise SME Knowledge Graphs. In *Proceedings of the 11th Knowledge*

*Capture Conference* (pp. 129-136). Association for Computing Machinery. https://doi.org/10.1145/3460210.3493569

Mo, F., Rehman, H. U., Elshafei, B., Chaplin, J. C., Sanderson, D., Martínez-Arellano, G. et al. (2023). Efficient Decision-Making in SMEs: Leveraging Knowledge Graphs with Neo4j and AI Vision. *IET Conference Proceedings, 2023,* 69-75. https://doi.org/10.1049/icp.2023.1736

Mukarram, S. S., Saleem, F., Alomair, A., & Al Naim, A. S. (2025). Knowledge Management and Sustainable Innovation for Green Finance: Evidence from Asia. *Humanities and Social Sciences Communications, 12,* Article No. 1742. https://doi.org/10.1057/s41599-025-06022-0

Park, Y.-J., Lee, M.-A., Yang, G.-J., Park, S. J., & Sohn, C.-B. (2022). Biomedical Text NER Tagging Tool with Web Interface for Generating Bert-Based Fine-Tuning Dataset. *Applied Sciences, 12,* Article No. 12012. https://doi.org/10.3390/app122312012

Pereira, N., & Fernandes, C. (2025). Knowledge Management in Health Organizations: A Systematic Literature Review. *Journal of the Knowledge Economy,* 1-73. https://doi.org/10.1007/s13132-025-02760-3

Pisoni, G., Molnár, B., & Tarcsi, Á. (2024). Knowledge Management and Data Analysis Techniques for Data-Driven Financial Companies. *Journal of the Knowledge Economy, 15,* 13374-13393. https://doi.org/10.1007/s13132-023-01607-z

Puspita, M. E., & Pramono, J. (2019). Factors Affecting the Use of Accounting Information in Small and Medium Enterprises (SMEs): A Study on SMEs in Tingkir, Salatiga. *The Indonesian Accounting Review, 9,* 207-225. https://doi.org/10.14414/tiar.v9i2.1818

Schulze, M., Pelzer, M., Schröder, M., Jilek, C., Maus, H., & Dengel, A. (2022). Towards Knowledge Graph Based Services in Accounting Use Cases. In *SEMANTICS 2022 EU: 18th International Conference on Semantic Systems* (pp. 1-5). CEUR Workshop Proceedings.

Sima, X., Coudert, T., Geneste, L., & de Valroger, A. (2022). Knowledge Management in SMEs: Preliminary Ideas for a Dedicated Framework. *IFAC-PapersOnLine, 55,* 1050-1055. https://doi.org/10.1016/j.ifacol.2022.09.528

Sutton, C., & McCallum, A. (2012). An Introduction to Conditional Random Fields. *Foundations and Trends® in Machine Learning, 4,* 267-373. https://doi.org/10.1561/2200000013

Theuma, A., Hayak, A. F., & Khdour, N. (2025). An Accounting Knowledge Management System That Fosters Innovation and Performance. *International Journal of Knowledge Management, 21,* 1-22. https://doi.org/10.4018/ijkm.365909

Wei, M., Xu, Z., & Hu, J. (2021). Entity Relationship Extraction Based on Bi-LSTM and Attention Mechanism. In *2021 2nd International Conference on Artificial Intelligence and Information Systems* (pp. 1-5). ACM. https://doi.org/10.1145/3469213.3470701

Wu, S., & He, Y. (2019). Enriching Pre-Trained Language Model with Entity Information for Relation Classification. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management* (pp. 2361-2364). ACM. https://doi.org/10.1145/3357384.3358119

Zhao, X. J., Blum, M., Yang, R., Yang, B. M., Carpintero, L. M., Pina-Navarro, M., Wang, T., Li, X., Li, H. T., Fu, Y. R., Wang, R. R., Zhang, J. T., & Li, I. (2024). *AGENTiGraph: An Interactive Knowledge Graph Platform for LLM-Based Chatbots Utilizing Private Data*.

Zhong, H., Chen, S. S., & Wang, M. (2024). Exploration of the Accounting Vertical Domain Reasoning Ability for Large Language Model. *Finance and Accounting Monthly, 45,* 17-25.

Zhou, P., Shi, W., Tian, J., Qi, Z., Li, B., Hao, H. et al. (2016). Attention-Based Bidirectional Long Short-Term Memory Networks for Relation Classification. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)* (pp. 207-212). Association for Computational Linguistics. https://doi.org/10.18653/v1/p16-2034

Zhu, J. H., Xia, Y. C., Wu, L. J., He, D., Qin, T., Zhou, W. G., Li, H. Q., & Liu, T. Y. (2020). Incorporating BERT into Neural Machine Translation. In *8th International Conference on Learning Representations (ICLR)*. (pp. 1-22). OpenReview.net.

## Appendix

**NER model-Detailed structure of BERT + CRF (See Figure 1):**

1) Input Processing Layer

- Input: Text sequence
- The model uses the tokenization method-WordPiece to convert the text into input_ids, attention_mask, and token_type_ids, and the sub-word sequence into token IDs for input into the BERT model after text input,
- The maximum text length is set to 256

2) BERT Encoding Layer

- Embedding layer: Word embedding, position embedding, and paragraph embedding
- Transformer layer: Multi-Head Attention Mechanism + Feedforward Neural Network
- Number of dimensions for hidden layers: 768 (base) or 1024 (large)
- Masked Language Modeling (MLM) layer: The entity label for each BERT output token is predicted, followed by prediction score calculation for each entity label, and the entity category with the highest probability for each token is assigned as its entity label

3) Output Layer

- Dimensional transformation: [batch_size, seq_len, hidden_size] → [batch_size, seq_len, num_labels]
- CRF layer: The optimal tag sequence are decoded after analysis of the transition probabilities between tags
- Output: The relevant labeled entity sequence is predicted

**NER model-Detailed structure of BERT + BiLSTM + CRF (See Figure 2):**

1) Input Processing Layer

Same as BERT + CRF

2) BERT Encoding Layer

Same as BERT + CRF

3) BiLSTM Layer

- Input layer: The representation of BERT output sequence [batch_size, seq_len, hidden_size]
- BiLSTM layer: Sequence information and long-range dependence are captured forward and backward
- Output layer: [batch_size, seq_len, lstm_hidden_size*2]
- Number of dimensions for hidden layers: 256 or 512

4) Output Layer

- CRF layer: Same as BERT + CRF
- Output: The relevant labeled entity sequence is predicted

**RE model-Detailed structure of BERT (See Figure 3):**

1) Input Processing Layer

◆ It is the same as the typical BERT input layer, but special tags are used to mark the entities, which provides the entities with the model the boundary and location information

- Special markers [E1] entity 1[/E1]... [E2] entity 2[/E2] are attached to the texts so as to identify the positions of the two entities
- The texts input to BERT are the ones with head and tail entity tags

2) BERT Encoding Layer

- Similar to the BERT encoding in the NER model, it obtains the contextual information of the entity pair and performs BERT encoding
- The input requires fields input_ids, segment_ids, and input_mask

3) Feature Extraction

- Both [CLS] and entity representation are used to mark representation
- [CLS] position embedding: The representation is marked with [CLS] as the sentence relation feature, and is added [CLS] and [SEP] at the beginning and end of each sentence respectively
- Embedding corresponding to two entities: Representations corresponding to a pair of entity tags are extracted, and spliced or average them to obtain the representation of the entity pair, learning the semantic features of the entity
- The three features obtained are spliced together and input into the following classification layer

4) Classification Layer

- Prediction of the relationship type between the head and tail entities
- Feed-Forward layer: mapping features to relation type space
- Softmax layer: the number of relationship types (including "no relationship") and the probability distribution of output for each relationship type

5) Output Layer

- Output: Entity-relation sets are generated, grouped and integrated by sentence to generate final prediction data

**RE model-Detailed structure of BERT (See Figure 4):**

1) Input Processing Layer

Same as BERT

2) BERT Encoding Layer

Same as BERT

3) BiLSTM Layer

- Further study of the relationship characteristics between entities
- Input: Sequence representation of BERT output
- Bidirectional LSTM: Long-range dependence in sequences are captured
- The embedded sequence is input into BiLSTM to get the concatenation vector of the forward and backward hidden states at each time step
- At each time step, BiLSTM outputs a hidden state sequence, and information at different positions contributes differently to the judgment of the relationship
- Output: [batch_size, seq_len, lstm_hidden_size*2]

4) Attention Layer

- Key semantics are extracted, focusing on the words or fragments most relevant to the relationship judgment

- An attention weight vector to weightedly sum the outputs of all time steps are learned, ultimately aggregating them into a vector that represents the global representation of the sentence
- Two entities in the sentence, divide the sentence into five parts, and perform pooling operation respectively, so that the model can learn the relevant features of the entities in the sentence
- Global maximum pooling: The most significant features are captured
- Global average pooling: Average features are captured
- The final feature representation is obtained from the splicing pooling results

5) Classification Layer

- Feed-Forward layer: features are mapped to relation type space
- Softmax layer: the number of relationship types (including "no relationship") and the probability distribution of output for each relationship type

6) Output Layer

- Output: Entity-relation sets are generated, grouped and integrated by sentence to generate the final predicted relation