UNIVERSITY OF TURKU

Identifying Supply Chain Risks using Large Language Models

Information and Communication Technology
Department of Computing, Faculty of Technology
Bachelor's thesis

Author:
Lasse Nordqvist

5.5.2025
Turku

The originality of this thesis has been checked in accordance with the University of Turku quality assurance system using the Turnitin Originality Check service.

Bachelor's thesis

Subject: Information and Communication Technology
Author: Lasse Nordqvist
Title: Identifying Supply Chain Risks using Large Language Models
Number of pages: 30 pages
Date: 5.5.2025

Supply chains (SC) are large and complex networks that consist of the flow of physical material, information, and assets between companies. The complexity of SCs has become a considerable problem in recent years, causing SCs to become vulnerable to a growing number of risks. The SC environment is also becoming more uncertain, as accurate data about the SC is scarce, and the number of disruptions in the SC is ever-growing. In the face of an uncertain SC environment, conventional SC risk management (SCRM) tools are becoming outdated.
Large language models (LLMs) have shown their ability to perform unformalized tasks, due to their exceptional natural language understanding and generation capabilities and specific domain knowledge acquired from a vast training dataset. LLMs have been successfully tailored and trialled for use cases in many academic fields, including SCRM.
LLMs can be tailored to identify SC risks (SCR) from textual data by utilizing prompt engineering and hyperparameter tuning. Natural language processing tasks, such as semantic text similarity analysis and named entity recognition, can be used to extract SCRs from textual data. LLM-based implementations are robust and proactive methods to support the conventional SCRM process in the face of the modern uncertain SC environment. SCRs identified by these implementations, however, lack the universal support of SCRM domain experts. To develop the SCR identification capabilities of LLMs, AI and SCRM domain experts should co-operate in creating a general taxonomy of SCRs.

Keywords: Large language model, supply chain risk management, risk identification, risk assessment

Table of contents

List of Figures

List of Tables

List of Abbreviations

Abbreviation	Definition	Abbreviation	Definition
AI	Artificial intelligence	LM	Language model
API	Application programming interface	LLM	Large language model
BERT	Bidirectional encoder representations from transformers	LR	Learning rate
BS	Batch size	MSL	Maximum sequence length
CAM	Causal attention mechanism	M&S	Marketing and Sales
CE	Contributing event	NA	News article
CR	Coreference resolution	NCAM	Non-causal attention mechanism
CRF	Conditional random field	NER	Named entity recognition
CS	Cosine similarity	NLG	Natural language generation
CTBR	Cambridge Taxonomy of Business Risks	NLP	Natural language processing
DOLLM	Decoder-only large language model	NLT	Natural language text
DR	Dropout rate	NLU	Natural language understanding
EDLLM	Encoder-decoder large language model	PE	Prompt engineering
EE	Event extraction	OM	Operations management
EOLLM	Encoder-only large language model	RA	Risk assesment
EV	Electric vehicle	RI	Risk identification
EVSC	Electric vehicle supply chain	RMI	Risk mitigation
FT	Fine-tuning	RMO	Risk monitoring
F&A	Finance and Accounting	R&D	Research and Development
GA	Genetic algorithm	SC	Supply chain
GCN	Graph convolutional network	SCE	Supply chain environment
GPT	Generative pre-trained transformers	SCM	Supply chain management
HCAM	Half-causal attention mechanism	SCR	Supply chain risk
HMLMM	Hybrid model large language model	SCRM	Supply chain risk management
HOR	House of Risk	SL	Schema learning
HR	Human resources	STSA	Semantic text similarity analysis
IE	Information extraction	TKL	Tokenized sequence length
IT	Information technology	TP	Temperature parameter
KPE	Key phrase extraction	WEV	Word embedding vector
LC	Logical constraint	WS	Warmup step

1 Introduction

The supply chain (SC) is a key concept for all manufacturing firms. An SC consists of all activities, such as production of finished goods, operators such as suppliers and retailers, and physical assets such as raw materials, that are needed to transform raw materials into finished products and services that are sold to customers. (Khan & Yu 2019, 4–6; Nakano 2019, 3–11.) Modern global SCs are becoming increasingly complex, spanning vast geographical areas and consisting of many tiers of suppliers and sub-supplier, with individual supplier tiers consisting of an increasing number of suppliers. As SCs are becoming increasingly complex, gathering accurate data on the SC is becoming more difficult, weakening a single company's ability to control of the operations within the SC. With added complexity, SCs are also becoming more vulnerable towards different kinds of supply chain risks (SCRs) that vary greatly in terms of the severity of disruption to the operational performance of SCs. (Aboutorab et al. 2022; Aboutorab et al. 2024; Aljohani 2023; Chu et al. 2020; Fiksel et al. 2015; Kassa et al. 2023; Kırılmaz & Erol 2017; Norrman & Jansson 2004; Vilko & Hallikas 2012.)

Due to the fast-paced and unpredictable nature of modern SC disruptions, conventional supply chain risk management (SCRM) methods are becoming outdated, as they are reactive in nature, rely heavily on historical data, and require accurate data on the SC to be useful. To cope with the uncertainties in the modern supply chain environment (SCE), companies need to adopt more proactive SCRM tools that are more robust to the scarcity of accurate SC information. (Aljohani 2023; Deiva Ganesh & Kalpana 2022; Kırılmaz & Erol 2017; Mukherjee et al. 2024; Zhao, M. et al. 2024a; Zhao, M. et al. 2024b.) One of these possible proactive SCRM tools is event and anomaly detection. When companies can proactively detect events that will possibly lead to disruptions in their SC, their ability to prepare mitigation strategies to these disruptions increases significantly. (Aboutorab et al. 2022; Aboutorab et al. 2023; Aboutorab et al. 2024; Aljohani 2023; Bodendorf & Zimmermann 2005.)

With the disruptions caused by SCRs materializing becoming more common, SCRM has become a subject of increasing scholarly interest, with the number of publications covering different SCRM themes increasing year by year (Ding & Huang 2024). One of the main themes in future SCRM research is combining SCRM with artificial intelligence (AI) and big data analytics (Aboutorab et al. 2022; Baryannis et al. 2018; Deiva Ganesh & Kalpana 2022;

Ding & Huang 2024; Kassa et al. 2023). Big data is characterized by the traditional three 'Vs': volume (consists of vast volumes of data), velocity (is created in real time), and variety (data can be structured, semi-structured or unstructured) (Kitchin & McArdle 2016). Unstructured textual data makes up the majority of useful data available worldwide (Ahammad et al. 2016; Sokolova 2018), a trend that is particularly significant in the realm of SCRM (Beheshti-Kashi et al. 2019; Sadeek & Hanaoka 2023; Shahsavari et al. 2024b). Given the substantial share of unstructured textual data in big data, large language models (LLMs) are crucial for big data analytics, due to their capacity to efficiently process, analyse, and extract valuable structured information from unstructured textual data (Roosan et al. 2024; Zhang, Y. et al. 2024). LLMs can be used, for example, in event and anomaly detection from big data (Cheng et al. 2024; Meng et al. 2024; Shahsavari et al. 2024b; Surampudi 2024, 1; Zhong et al. 2025), meaning they have considerable potential to be used SCR identification process.

This thesis aims to shed light on how LLMs can be used to identify SCRs and how LLM-based risk identification can support the conventional SCRM process. The thesis aims to answer the following research questions:

RQ1: How can LLMs be used to detect SCRs?

RQ2: What kind of SCRs can be reliably extracted using LLMs?

RQ3: How can the use of LLMs support the conventional supply chain risk management process?

RQ4: How can the use of LLMs in identifying supply chain risks be improved in the future?

This thesis has been conducted as a literature review with a small empirical section in section 4.4. Articles for this thesis were searched using Google Scholar and the IEEE Xplore database with the search statements ("llm" OR "large language model") AND ("supply chain risk" OR "SCRM") and "supply chain" AND "Cambridge taxonomy". Using the former search statement, Google Scholar returned 550 articles, and the IEEE Xplore database returned 43 articles. Using the latter search statement, Google Scholar returned 17 articles, and the IEEE Xplore database returned 2 articles. To supplement the articles found by using Google Scholar and the IEEE Xplore database, articles were also searched by using SCOPUS AI and Consensus AI search engines for academic publications. For the AI search engines, the prompts used to search relevant articles were "large language models in supply chain risk management" and "using large language models to identify supply chain risks". The articles returned by these search methods were filtered based on the relevance of the article's title, abstract, and contents in terms of the thesis's research subject, and date of publication. Articles published after 2017 were taken into consideration to be used in the thesis, based on the landmark paper by Vaswani et al. (2017), that introduced LLMs to mainstream academic research. By utilizing these filtering parameters, 10 articles were identified as relevant articles to the thesis.

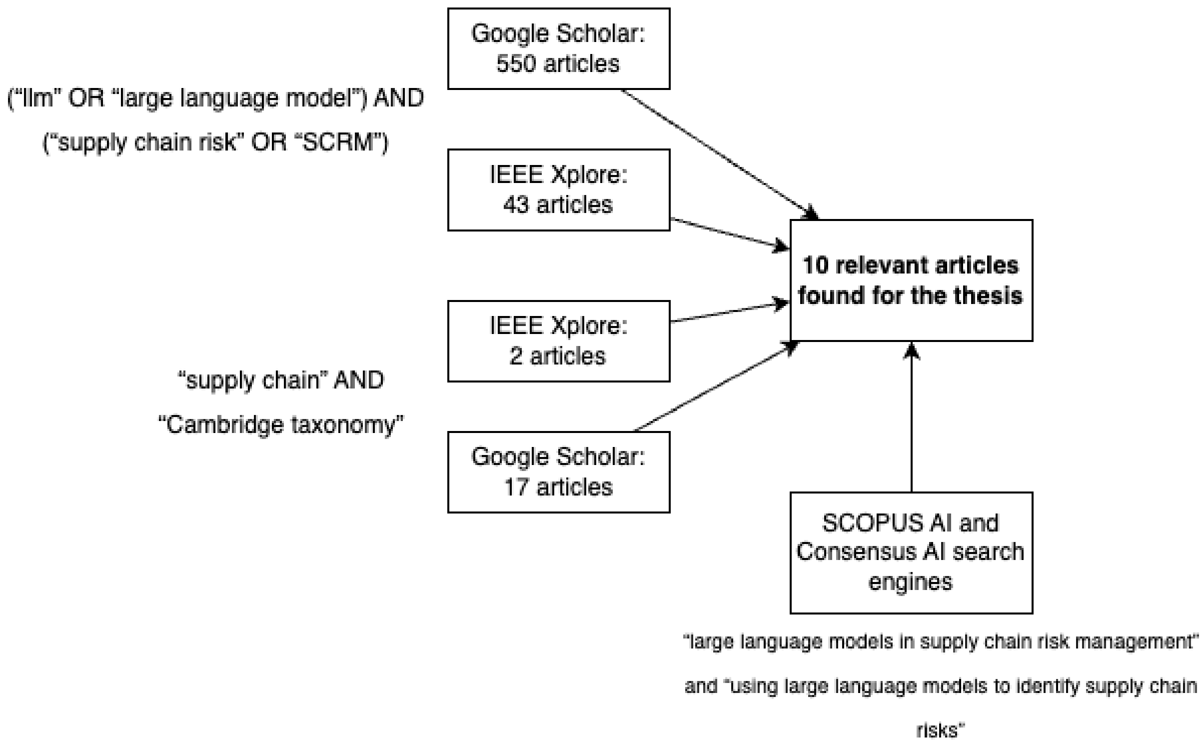

Figure 1. Results for the used search methodologies

There was considerable overlap between the different search methods, as most articles were included in the results of multiple search methods. The articles that were found using the above-mentioned methodologies are introduced in more detail in Table 2.

2 Supply chain management

Supply chain management (SCM) can be defined as “strategic and organizational management that enables a focal firm to overcome performance trade-offs and realize operational competitiveness through a fit among the management elements within and across the firm to adapt to the external environment” (Nakano 2019, 19). SCM is a very ambiguous concept that requires functioning cross-organizational co-operation. SCM can be viewed as a management philosophy that views the SC as a whole and complex system. SCM aims to foster the operational and strategic capabilities of the different firms collaborating in the SC to create additional value in the SC and to optimize the flow of goods inside the SC. To realize the benefits of the created SCM philosophy, companies need to establish suitable management practices, activities, and processes to implement the created philosophy in practice. (Mentzer et al. 2001; Mentzer et al. 2008.)

2.1 How different business functions engage in SCM

Within organizations, the structure of SCM remains ambiguous. SCM is not “owned” by any business function (e.g. in a way that accounting handles accounting), but is instead a phenomenon that is interconnected to nearly every business function. (Mentzer et al. 2008.) Due to the cross-functional nature of SCM, effective cross-functional co-operation is essential for its success (Badwan 2024).

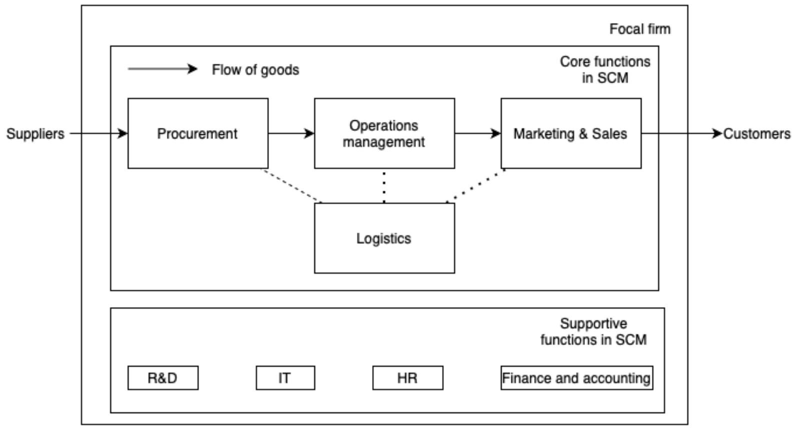

Figure 2. SCM process in companies, figure adapted from (Helmold & Samara 2019, 51; Lambert 2003, 3; Mentzer et al. 2008; Nakano 2019, 20)

Different business functions engage in SCM in different ways. In SCM literature, the following four functions are usually mentioned as core business functions that engage in SCM:

Procurement is responsible for the acquisition of the supply-side goods and services that are needed in the production of finished goods and services (Helmold & Samara 2019, 51–52; Khan & Yu 2019, 191). Procurement is responsible for many operational and strategic activities that are related to SCM. Procurement handles supply management and manages all the suppliers that the company supplies goods and services from. Supply management is a key strategic process in SCM that requires continuous supplier performance monitoring. (Helmold & Samara 2019, 51–65; Khan & Yu 2019, 191–205.) On top of strategic duties, procurement is also responsible for the operational acquisition of goods and services needed by the company. The overall goal of the procurement function is to ensure that all goods and services needed by the company are acquired in a timely and cost-efficient manner. (Khan & Yu 2019, 192–193.)

Operations management (OM) is responsible for the design and controlling of the company's production system that converts inputs (e.g., raw material, labor) into outputs (e.g. goods or services). OM controls this flow through different strategic and operational activities, that also affect the implementation of SCM. (Helmold & Samara 2019, 67–68; Mentzer et al. 2008.) In terms of SCM engagement, OM can, for example, require that procurement acquires certain inputs at a certain price and that logistics transports these inputs to OM's disposal at a certain time. The primary objective of OM is to ensure the efficiency of the company's operations, maximize product quality, and minimize resource usage during the production process (Helmold & Samara 2019, 67–68).

Marketing and Sales (M&S) is responsible for creating, communicating, and delivering the value of the company's outputs to customers in a way that is beneficial to the company and its shareholders. M&S defines the long-term strategic value creation objectives and executes these objectives on the operational front through sales work and customer relationship management. (Guenzi & Troilo 2007; Mentzer et al. 2008.) In the context of SCM engagement, M&S may, for example, launch a marketing campaign. This campaign requires OM to produce enough outputs. Logistics must also have the capacity to distribute the outputs to customers. This ensures the company can meet the increased demand resulting from the campaign.

Logistics is responsible for transporting and storing all the needed materials flows that are present in Figure 2 and defines how these material flows will be executed. Logistics aims is to find an optimized SC configuration for material flows inside the SC, by defining, for example, storage locations, inventory levels, distribution center locations, and used transportation methods (Mentzer et al. 2008; Nakano 2019, 3; Wang 2012). SCM literature also identifies some business functions that take a more supportive role in engaging in SCM:

Research and Development (R&D) is responsible for creating new and improving the company's existing products and processes (Forbes & Wield 2004). R&D is known for demanding high-quality inputs for products it develops. R&D often emphasizes that the company should source from suppliers that can supply high quality inputs for a reasonable price (Xie et al. 2016). R&D can also increase a supplier's strategic importance in the company's SC strategy by involving and collaborating with suppliers in the development of new products (Delgado-Verde & Díez-Vial 2024).

Information Technology (IT) is responsible for implementing and maintaining information systems that all other business functions use (Dewett & Jones 2001; Zand et al. 2015). IT has a crucial role in supporting the information flows between different operators of the SC (Varma & Khan 2014) that are crucial in the overall SCM process (Lambert 2008, 3). IT supports the information flow by implementing and maintaining technologies, such as the enterprise resource planning system, radio frequency identification, and electronic data interchange to be used in the SC (Varma & Khan 2014).

Human resources (HR) is responsible for the management and development of the company's workforce (Helmold & Samara 2019, 155). Given the complex nature of SCs, HR has a crucial role in fostering SCM knowledge in companies and enhancing SC performance (Gómez-Cedeño et al. 2015; Haq et al. 2021).

Finance and Accounting (F&A) is responsible for the financial reporting, planning, and management of companies (Chang et al. 2014). F&A has a role in aligning the financial flows with information and material flows within SCs, thereby improving the company's cash flow management (Xu et al. 2018). Moreover, F&A can assist in establishing SCM cost objectives (Almatarneh et al. 2022).

2.2 Supply chain risk management

The conventional SCRM framework is largely based on the conventional business risk management framework (Blos et al. 2009). Conventional business risk management tools can be applied in SCRM. Failure Mode and Effects Analysis, that aims to identify potential failures in business systems and their causes (Nyoman Pujawan & Geraldin 2009), and the

ISO 31000 business risk management process standard (de Oliveira et al. 2017; Hermoso-Orzáez & Garzón-Moreno 2021) are examples of conventional business risk management tools.

2.2.1 Conventional SCRM process

SCRM literature usually defines the conventional SCRM process as a four-stage circular process (Ho et al. 2015; Hermoso-Orzáez & Garzón-Moreno 2022; Norrman & Jansson 2004; Olson 2014, 13–15; Pham et al. 2023; Rangel et al. 2015 Tummala & Schoenherr 2011).

Risk identification (RI) is the first stage in the SCRM process, in which the aim is to identify the risks that are relevant to the SC and the factors that could cause these risks. SCRs are often categorized by some aggregate factor (e.g., demand risk, supply risk). Most often, operational SCRs are identified using quantitative methods, while more long-term SCRs are identified using qualitative analysis. (Ho et al. 2015; Hermoso-Orzáez & Garzón-Moreno 2022; Norrman & Jansson 2004; Olson et al. 2014, 13; Pham et al. 2023; Rangel et al. 2015; Tummala & Schoenherr 2011.)

Risk assessment (RA) succeeds the RI stage. In the RA stage, the aim is to assess the probability of occurrence and impact on the SC for every SCR identified in the RI stage. Occurrence probabilities and impact assessments are used to prioritize identified SCRs; a high occurrence probability and effect on the SC indicate a high priority SCR that the company should focus on, and vice-versa. Different categories of SCRs have different kinds of assessment methods that are most suited to them. (Ho et al. 2015; Hermoso-Orzáez & Garzón-Moreno 2022; Norrman & Jansson 2004; Olson et al. 2014, 13–14; Pham et al. 2023; Rangel et al. 2015; Tummala & Schoenherr 2011.)

Risk mitigation (RMI) succeeds the RA stage. In the RMI stage, companies turn their attention to the high-priority SCRs defined in the RA stage and try to define mitigation strategies for these SCRs and carry out these strategies at the operational level. Companies should focus on the high-priority risks because defining functional mitigation strategies is a time and cost-consuming process. SC configuration redesigns are often used to mitigate SCRs. If, for example, there is a risk that a critical component for a company's operations might run out, the company might opt to stockpile these components to ensure supply even during SC disruptions. Similar to the assessment of SCRs, different mitigation strategies need to be tailored for each SCR. (Ho et al. 2015; Hermoso-Orzáez & Garzón-Moreno 2022;

Norrman & Jansson 2004; Olson et al. 2014, 14–15; Pham et al. 2023; Rangel et al. 2015; Tummala & Schoenherr 2011.)

Risk monitoring (RMO) is the final stage of the conventional SCRM process and succeeds the RMI stage. In the RMO stage, the identified SCRs should be monitored and reviewed continuously to see how implemented mitigation strategies have affected the SCRs and to identify possible shift in the SCE and SCRs. If SCRM monitoring suggests that the SCRs are changing, the company needs to restart the SCRM process and return to the RI stage and try to identify the new emerging SCRs, completing the circular process of SCRM. (Ho et al. 2015; Hermoso-Orzáez & Garzón-Moreno 2022; Pham et al. 2023; Rangel et al. 2015; Tummala & Schoenherr 2011.)

One example of a systematic framework for the SCRM process is the House of Risk method (HOR) introduced by Nyoman Pujawan & Geraldin (2009). The framework is based on identifying SCR risk events and factors that contribute to these events, assessing what impact these events would have on the supply chain, and prioritizing mitigation strategies based on their effectiveness and their criticality.

Figure 3. Conventional SCRM process, figure adapted from (Ho et al. 2015; Hermoso-Orzáez & Garzón-Moreno 2022; Norrman & Jansson 2004; Olson 2014, 13–15; Pham et al. 2023; Rangel et al. 2015; Tummala & Schoenherr 2011)

2.2.2 SCR classification

SCRs need to be classified because methods used in all SCRM stages are tailored to specific types of SCRs (Ho et al. 2015). In SCRM literature, SCRs have been classified in various ways and dimensions. Commonly, SCRM literature defines two aggregate-dimension classifications of SCRs.

Relation to SC defines how the factors causing the SCR relate to the SC network. SCRs can be caused by factors that are external to the SC network (e.g., climate change), internal to the SC network but external to the focal firm (e.g., supplier performance) or internal to the focal firm (e.g., production machine breakdown). (Jüttner et al. 2003; Norrman & Jansson 2004; Pandey et al. 2023; Shahbaz et al. 2019; Vilko & Hallikas 2012.)

SC dimension defines where in the SC the impact of the SCR is most prominent. SCRs can be related to the supply side (e.g., supplier performance), operational side (e.g. manufacturing) or demand side (e.g., customer demand). (Fiksel et al. 2015; Kleindorfer & Saad 2005; Norrman & Jansson 2004; Pandey et al. 2023; Shahbaz et al. 2019; Vilko & Hallikas 2012.)

These aggregate-level classifications can be differentiated further. Table 1 outlines various types of SCRs, their connection to the aggregate level classifications, and provides examples of their causes and potential effects on the SC.

Table 1. Types of SCRs

SCR type	SCR causes	SCR effects	Factor	Primary effect SC	References
Political risk	Terrorism, war, tariffs, poor public policy	Supply disruption, increase in cost of materials	External	Supply	Chu et al.; Fiksel et al.; Jüttner et al.; Kleindorfer & Saad; Pandey et al.; Rangel et al.; Shahbaz et al.
Environmental risk	Climate change, natural disasters, industrial disaster, pandemic	Supply disruptions	External	Supply	Chu et al.; Fiksel et al.; Jüttner et al.; Kleindorfer & Saad; Norrman & Jansson; Rangel et al.; Shahbaz et al; Vilko & Hallikas
Financial risk	Lack of investments, poor financial performance	Suppliers withdrawing deliveries due to unpaid invoices, unfavorable terms of payment	Operational	Supply, Operational	Chu et al.; Ho et al.; Pandey et al.; Pham et al.; Rangel et al.; Shahbaz et al.
Supply risk	Loss of supplier, poor supplier performance	Supply disruption, poor quality of materials	Internal	Supply, Operational	Chu et al.; Ho et al.; Kleindorfer & Saad; Pandey et al.; Pham et al.; Rangel et al.; Shahbaz et al.
Demand risk	Fluctuating demand, customer fragmentation	Difficulties in SCM forecasting, need for supplier flexibility	Internal	All	Chu et al.; Ho et al.; Norrman & Jansson; Pham et al.; Pandey et al.; Rangel et al.; Shahbaz et al.
Logistics risk	Inefficient use of transport capacity, cargo damage	Increase in transportation costs, contract fines	External	Supply, Demand	Chu et al.; Ho et al.; Pham et al.; Rangel et al.; Shahbaz et al.; Vilko & Hallikas
System risk	Technical debt in the SC, over-dependency on certain suppliers	Inefficient SCM processes, SC vulnerable to critical supplier loss	Internal, Operational	All	Chu et al.; Fiksel et al.; Pandey et al.; Rangel et al.
Operational risk	Employee strike, machine breakdown	Halt in production, demand not met	Operational	Operational, Demand	Chu et al.; Fiksel et al.; Ho et al.; Jüttner et al.; Pandey et al.; Rangel et al.; Shahbaz et al.
Social risk	Employee exploitation in the SC, lack of SC transparency	Reputational damages, fines	Internal	Supply, Demand	Pandey et al.
Cyber security risk	Cyberattacks, inadequate data handling practices	Data breaches, reputational damage	Internal, Operational	Supply, Operational	Pandey et al.
Behavioural risk	Unbalanced information exchange	Lack of trust towards suppliers, bullwhip effect	Internal, Operational	Supply, Operational	Pandey et al.; Rangel et al.; Shahbaz et al.
Cultural risk	Lack of knowledge about cultural differences in global SC	SCM process inefficiencies, conflicting interpretations of contracts	Internal, External	Supply, Demand	Fiksel et al.; Rangel et al.

3 Large language models

The history of language models (LM) dates back to the 1950s and 60s, when the first LMs were introduced and performed simple natural language processing (NLP) tasks such as binary classification. During the 1980s and 90s, statistics-based LMs were introduced that extracted statistical patterns from text. Mid-2000s saw a key breakthrough in NLP technology when word embedding vectors (WEV), vector-space representations of words and text, were first introduced. WEVs have a key role in LLMs. The 2010s saw significant breakthroughs in LMs, neural networks were first introduced in NLP tasks, and in 2017, transformers were introduced. The introduction of transformers can be seen as a turning point in LLMs, as most modern state-of-the-art LLMs are based on transformer technology. (Raiaan et al 2024.)

3.1 Definition of LLMs

LLMs are a category of AI models that are developed to comprehend, analyze, and produce human language. LLMs are characterized by extensive training on vast datasets, the massive number of parameters the model is trained on (up to multiple trillions), and the self-supervised learning training approach. LLMs are built upon the theory of NLP, a field of machine learning that aims extract knowledge from natural language text (NLT) and allow computers to achieve natural language understanding (NLU) and use the NLU for natural language generation (NLG). (Omar et al. 2024; Raiaan et al. 2024; Singh & Mahmood 2021; Yang et al. 2024; Yao et al. 2024.) Yang et al. (2024) list five key features that modern LLMs should possess, as follows:

Deep understanding of NLT: LLMs should have a deep understanding of NLT so they can extract meaning from it and perform NLP tasks based on the extracted meaning.

Generation of human-like text: LLMs should be able to generate human-like text based on the extracted meaning from NLT and the prompt it has been given.

Contextual awareness: LLMs should be able to utilize their vast training data from different contexts and demonstrate domain expertise. LLMs should be able to adjust their responses if prompts want them to show expertise in a specific domain (e.g., ‘What kind of supply chain risks are there from a reputational perspective?’).

Excellence in problem-solving and decision-making: LLMs should show excellence in problem-solving and decision-making tasks by leveraging their vast training data.

Applying abilities in real-world scenarios: LLMs should be able to complete unformalized tasks given to them by harnessing the four aforementioned abilities.

3.1.1 Main types of modern LLM architectures

Transformers introduced by Vaswani et al. (2017) represented a breakthrough in LLMs, and modern LLMs are largely based on transformers. Yang et al. (2024) define the following four main types of modern LLM architectures, as follows:

Encoder-only LLMs (EOLLM) are used for NLU tasks and analyzing and extracting knowledge from input data. EOLLMs focus on capturing contextual information in input sequences and are computationally cheaper than the other LLMs, reaching inferences faster. EOLLMs require less training than the other LLM architectures. EOLLMs are used with prompts; the LLM analyzes the input data and generates a specific kind of analysis based on the prompts given to it. EOLLMs can be used for tasks such as sentiment analysis, document classification, and feature extraction. Bidirectional encoder representations from transformers (BERT) are the most well-known examples of EOLLMs (Benayas et al. 2024; Soliman et al. 2025.)

Decoder-only LLMs (DOLLM) were originally developed for NLG tasks, but state-of-the-art DOLLMs also excel in different kinds of NLU tasks and can be used for the same tasks as EOLLMs. The extensive pre-training of DOLLMs enables them to learn domain-specific knowledge, helping in NLU tasks. DOLLMs receive a prompt and generate an answer to the prompt based on the pre-training. Modern DOLLMs can understand prompts and generate answers in different kinds of forms, ranging from text, image, audio, and video. The problem with DOLLMs is the need for considerable computational resources with the extensive pre-training of complex models with billions of parameters. Generative pre-trained transformers (GPT) are the most well-known examples of DOLLMs. (Benayas et al. 2024; Caillaut et al. 2024; Soliman et al. 2025; Yang et al. 2024.)

Encoder-decoder LLMs (EDLLM) are similar to EOLLMs, but the models also include a decoder. EDLLMs represent the original architecture of a transformer, containing both an encoder and a decoder. Like EOLLMs, EDLLMs focus on the input sequence and create a compressed representation of the input sequence. The decoder is then used to generate an output sequence based on the encoder's results, most often in text-form. Because of the dual-architecture, EDLLMs are limited in their ability to support certain kinds of tasks and lag behind the performance levels of EOLLMs and DOLLMs that can be better adapted to all kinds of tasks. (Livne et al. 2024; Raiaan et al. 2024; Wang et al. 2023; Yang et al. 2024.)

Hybrid model LLMs (HMLLM) combine the traditional EDLLMs with the pre-training of DOLLMs. The input sequences given to the model are a combination of EOLLM and EDLLM input sequences. The introduction of pre-training is used to enhance the generative capabilities of the model. (Yang et al. 2024.)

3.1.2 Main types of modern attention mechanisms in LLMs

Attention mechanisms are a key concept in LLMs. Attention mechanisms affect how the LLM handles input sequences and how it creates links between tokens inside the input sequence. (Raiaan et al. 2024; Vaswani et al. 2017; Yang et al. 2024.) Yang et al. (2024) define the three types of attention mechanisms used in modern LLMs, as follows:

Causal attention mechanisms (CAM) allow the LLM to access only past tokens inside the input sequence when making predictions. LLMs with CAMs can reach relatively low latency and high inference speeds, which makes them suitable for real-time applications that require quick response times. CAMs are most often used in DOLLMs. (Strimel et al. 2023; Yang et al. 2024.)

Non-causal attention mechanisms (NCAM) allow the LLM to access all tokens inside the input sequence. NCAMs make predictions based on the context built by the past and future tokens. LLMs with NCAMs reach better predictive accuracy compared to CAMs due to the attention mechanism having the ability to access the whole input sequence. However, accuracy gains come with increased computational costs, increased latency, and slower inference speeds. NCAMs are most often used in EOLLMs. (Strimel et al. 2023; Yamaguchi et al. 2023; Yang et al. 2024.)

Half-causal attention mechanisms (HCAM) combine the causal and non-causal attention mechanisms and chooses the appropriate mechanism based on the task at hand and its need for precise accuracy or fast latency (Strimel et al. 2023; Yang et al. 2024).

3.2 Use of LLMs in an SCM context

LLMs have vast use cases in the context of SCM. As an antecedent to using LLMs in an SCM context, companies need to have functioning data and knowledge management processes, so that insights extracted by LLMs can be easily shared with decision-makers. LLMs also require domain-specific data on SCs that they can draw insights from. (Aghaei et al. 2025; Zhang, J. et al. 2024.) Zhang, J. et al. (2024) have identified four emerging use cases for LLMs in a SCM context, as follows:

Contract management handles the different contracts related to SCs, such as supplier and customer contracts. LLMs can be used streamline the contract management process and help in the extraction of contract intelligence that can be used to negotiate better contracts for the company. (Aghaei et al. 2025; Bellomi & Cristani 2024; Vlachos & Pulagam 2025; Zhang, J. et al. 2024.) For example, Ito & Nakagawa (2024) introduced an LLM-based tool that analyses and helps decision-makers understand public tender documents, that outline contract terms for companies participating in public procurement.

SCM knowledge management centers around the extraction, analysis, storage, and sharing of different SCM knowledge to support SCM decision-making. LLMs can extract insights from, for example historical sales data, news articles (NA) or economic indicators. The responsibility of LLMs is to make this knowledge accessible to decision-makers. (Aghaei et al. 2025; Vlachos & Pulagam 2025; Zhang, J. et al. 2024.) For example, Gezdur & Bhattacharjya (2025) introduced an LLM-based chatbot for training SC employees and assisting them in everyday SC work.

SCRM is one of the main use cases of LLMs in the SCM context. LLMs can be used to detect SCRs proactively, helping to form risk mitigation strategies. (Aghaei et al. 2025; Vlachos & Pulagam 2025; Zhang, J. et al. 2024.) Numerous articles have been published on LLMs identifying and labeling SCRs (Kühl et al. 2024; Shahsavari et al. 2024a; Shahsavari et al. 2024b; Zhao, M. et al. 2024a; Zhao, M. et al. 2024b).

SC optimization aims to optimize the SC configuration by changing, for example, inventory strategies. By utilizing LLMs in SC optimization, companies can achieve improved SC performance and decreased SC costs. (Aghaei et al. 2025; Vlachos & Pulagam 2025; Zhang, J. et al. 2024.) For example, Li et al. (2024) introduced an LLM-based system for optimizing inventory management in SCs.

4 LLMs in extraction of SCRs

LLMs have shown potential that they can be used to extract and identify SCRs. All of the articles mentioned in Table 2 are based on the premise of utilizing LLMs for information extraction (IE), the process of extracting structured knowledge from unstructured data, in which LLMs are very adept (Roosan et al. 2024; Zhang, Y. et al. 2024). There are some prerequisites before companies can utilize LLMs in extracting SCRs. Focal companies must have organized supplier databases where supplier information is easily accessible. Supplier information is needed so that LLMs can be tailored to extract SCRs that are related to the focal company's suppliers. (Zhao, M. et al. 2024a; Zhao, M. et al. 2024b.) Most of the articles mentioned in Table 2 extract SCRs from NAs, and most often these implementations include a web crawler collecting NAs that include information about SCRs. The selection of news outlets where NAs are collected from is crucial to ensure the journalistic integrity of sources, especially in implementations that include a geographical frame for the NAs. The dataset should also be as rounded as possible, global news should originate from different geographical locations, NAs should be collected from news outlets with different political orientations, and they should cover different topics (e.g., politics, economics), to align with the different risks introduced in Table 1. (Shishehgarkhaneh et al. 2024a.)

Table 2. Articles introducing implementations for SCR identification utilizing LLMs

Author & year	Title	Publication type	LLM architectures utilized	SCRM processes supported	Brief description of implementation
Cheng et al. (2024)	*SHIELD: LLM-Driven Schema Induction for Predictive Analytics in EV Battery Supply Chain Disruptions*	Conference	EOLLM, DOLLM	All	Pipeline that aims to extract disruption events contributing to SCRs in electric vehicle SCs from NAs. Assessment of extracted events based on SCR event schema tailored for the electric vehicle SC.
Fan et al. (2025)	*Measuring firm-level supply chain risk using a generative large language model*	Journal	DOLLM	RI, RA	DOLLM-based implementation that assesses company's exposure to SCRs, based on the analysis of investor site analysis transcriptions
Kühl et al. (2024)	*Enhancing Supply Chain Risk Identification: Analyzing the Impact of LLM Parameters for precise Classification*	Conference	DOLLM	RI	DOLLM-based implementation that examined the optimal temperature parameter value for DOLLMs tasked with categorizing SCRs.
Shahsavari et al. (2024a)	*Empowering Supply chains Resilience: LLMs-Powered BN for Proactive Supply Chain Risk Identification*	Conference	EOLLM, DOLLM	All	Pipeline with six modules to extract events contributing to SCRs from NAs to create a sequence of events contributing to a SCR.
Shahsavari et al. (2024b)	*Event Identification for Supply Chain Risk Management Through News Analysis by Using Large Language Models*	Journal	EOLLM, DOLLM	All	Based on the earlier article by Shahsavari et al. (2024b). Introduces added examples of use.
Shishehgarkhaneh et al. (2024a)	*Transformer-based Named Entity Recognition in Construction Supply Chain Risk Management in Australia*	Journal	EOLLM	RI	EOLLM-based implementation that explored the fine-tuning of model hyperparameters to perform NER to extract SCRs from NAs.
Shishehgarkhaneh et al. (2024b)	*Named Entity Recognition in Construction Supply Chain Risk Management using Transformer-Based Models and Genetic Algorithm-Driven Hyperparameter Tuning*	Conference	EOLLM	RI	Based on the earlier article by Shishehgarkhaneh et al. (2024a). Introduces a genetic algorithm for hyperparameter tuning.
Sun et al. (2024)	*Application of News Analysis Based on Large Language Models in Supply Chain Risk Prediction*	Journal	DOLLM	RI, RA	DOLLM-based implementation to extract SCRs from NAs and predicting future events based on extracted risks.
Zhao, M. et al. (2024a)	*Optimizing Supply Chain Risk Management: An Integrated Framework Leveraging Large Language Models*	Conference	EOLLM, DOLLM	RI	Pipeline that first extracts SCRs from NAs using a DOLLM and uses a EOLLM to categorize the extracted risks.
Zhao, M. et al. (2024b)	*Enhancing Supply Chain Risk Management with Large Language Models: Software Prototyping and Interactive Visualization*	Conference	EOLLM, DOLLM	All	Based on the earlier article by Zhao, M. et al. (2024a). Develops the pipeline further and introduces a full pipeline that supports all SCRM stages.

4.1 Engineering LLMs to extract SCRs

4.1.1 Prompt engineering

Prompt engineering (PE) has been used in several of the articles utilizing DOLLMs. Prompts are used to control LLMs and the outputs they produce. PE aims to develop prompts that increase the performance of LLMs in various tasks. By utilizing PE, LLMs can be tailored to solve intricate tasks that require domain-specific expertise, such as extracting SCRs from NAs. Prompts used in PE usually include task description, contextual settings, a question the LLM must solve, and examples or data related to the task. (Hussain et al. 2023; Polat et al. 2025; Wang et al. 2024; Zhao, M. et al. 2024a.)

Fan et al. (2025) utilized a GPT-4o mini DOLLM to measure the company's exposure to SCRs from the transcriptions of corporate site visit audits. PE was used to define different SCR categories to the model, and the model was tasked to identify these SCR categories from audit transcriptions.

Shahsavari et al. (2024a) and Shahsavari et al. (2024b) utilized a GPT-3.5 DOLLM to extract information about events contributing to SCRs from relevant NAs that were given as input. Utilizing PE, the model was asked to answer where the event happened, when it took place, and what was the duration of the contributing event (CE).

Zhao, M. et al. (2024a) also utilized a GPT-3.5 DOLLM and gave it NAs as input. They used PE to make the LLM take the role of a risk assessor at a focal company and asked the model to identify if the NAs given as input were related to the focal company and its suppliers. If relations were found, the model was asked to identify SCRs in the articles and generate a text output of the identified SCRs using PE.

Zhao, M. et al. (2024b) developed the system further utilizing PE, using the GPT-3.5 DOLLM to also define the likelihood of the SCR happening and the impact it would have on the SC.

4.1.2 Hyperparameter tuning

The temperature parameter (TP) controls the randomness of output generated by generative DOLLMs. Tuning the TP modifies the probability distribution of generated outputs. Lower TP values lead to more concentrated word generation probabilities, resulting in deterministic and precise outputs. Higher TP values lead to uniform word generation probabilities, resulting in more random outputs. The change in distribution is related to the softmax function that is commonly used in GPT LLMs to convert the raw scores of WEVs into a probability distribution. The basic formula of the softmax function can be stated as $P_i = \frac{e^{\frac{z_i}{\varepsilon}}}{\sum_j e^{\frac{z_j}{\varepsilon}}}$, where P_i is the probability that the next word in the generated output sequence will be the ith member of the WEV, Z_i is the raw word embedding score for the ith member, and ε the TP. Changing the value ε modulates the scaling of the exponential term, thus affecting the calculated probability distribution. Lower TP values are suitable to use when LLMs are prompted with factual tasks that have homogenous and precise answers defined for them, while higher TP values are suitable to use when LLMs are prompted with more creative questions that have ambiguous answer possibilities. The temperature parameter can be often modified by accessing the LLMs application programming interface (API). (Davis et al. 2024; Grandi et al. 2025; Qiu et al. 2024; Rosoł et al. 2023; Wang et al. 2024; Yan et al. 2024.)

Kühl et al. (2024) examined how the value of the TP value affected the LLMs' ability to classify SCRs. They defined that with a value interval of 0-2, a TP value between 0.4 and 0.7 produced the best classifying results. With a higher TP value, the LLM generated incomprehensible outputs that did not include any classification. They argued that a lower TP value would, however, lead to the LLM disregarding the complex interdependencies between the SCRs and their classification. Overall, it was hard to define a clear optimal TP value because the classification accuracy of the LLM was relatively stable for TP values in the range of 0-1.1. This minor effect of the TP is in line with other TP tuning research results when LLMs are prompted with tasks that require the comprehension of complex interdependencies. For example, Grandi et al. (2025) discovered that TP tuning does not have much effect on an LLM's ability to conduct material selection for engineering projects. Project engineering often has strict requirements for materials used in the project and material selection is a task that assesses how different materials fulfill these requirements. The interdependencies between the requirements and features make material selection a complex task.

Fine-tuning (FT) is the process of tailoring an LLM to a specific task. Due to their smaller complexity, EOLLMs can be better fine-tuned than DOLLMs and outperform them in most tasks where FT is applied, making them a viable and computationally cheaper option to perform tasks that have been narrowed down to a specific context. FT of DOLLMs is developing however, and they can already match the performance of EOLLMs in certain FT applications. (Benayas et al. 2024; Feroze et al. 2025; Huang et al. 2025.)

FT of EOLLMs includes applying conventional machine learning methods, such as train-test-validation splitting and gradient descent, to optimize the hyperparameters of pre-trained EOLLMs. The following are some key hyperparameters of EOLLMs. **Batch size** (BS) controls the number of samples that are used in gradient calculations that update the parameter weights in the training process. A higher BS will lead to faster training times but might reduce the prediction accuracy in the process. Selecting an appropriate BS is key in balancing the trade-off between faster training times and prediction accuracy. (Shishehgarkhaneh et al. 2024a; Zhao, L. et al. 2024.) **Learning rate** (LR) controls the step size of the model's parameter updates towards the gradient direction. A small LR means that updates are smaller and more precise. The downside of a small LR is that it might lead to slower convergence to the optimum and lead to the model getting stuck in a local optimum. A high LR means that updates are greater but having a high LR might lead to the model oscillating around the optimum, preventing convergence to the optimum. (Shishehgarkhaneh et al. 2024a; Zhao, L. et al. 2024.) **Epochs** define how many times the model is trained on the whole training set. A small number of epochs can result in underfitting, while a large number can result in overfitting. (Shishehgarkhaneh et al. 2024a; Zhao, L. et al. 2024.) **Maximum sequence length** (MSL) controls the tokenized sequence lengths (TKL) of inputs given to the model. The training of EOLLMs requires the TKLs of training samples to be fixed. If the sample's TKL is greater than the MSL, the training sample is truncated and part of it is cut off. If the TKL is smaller than the MSL, the sequence is padded with content related to the sequence that varies based on the chosen padding methods. Increasing the MSL increases the model's prediction accuracy but at the cost of also increasing the model's computational complexity. Choosing the right MSL is key in balancing the trade-off between predictive accuracy and computational complexity. (Ang et al. 2022; Dang et al. 2024; Dukić & Šnajder 2024; Shishehgarkhaneh et al. 2024a.) **Dropout rate** (DR) controls the activation and deactivation of neurons in the pre-training and FT of EOLLMs. DR is used to allow the model to learn more robust representations of the training data to improve the generalization ability of the model and avoid overfitting to the training data. A too high DR will however create problems with inconsistencies between training and future inferences of the model. (Ni et al. 2024; Shishehgarkhaneh et al. 2024a.)

Shishehgarkhaneh et al. (2024a) fine-tuned various EOLLMs for named entity recognition to extract SCRs from NAs. The EOLLMs were fine-tuned by using an annotated training set that contained SCR entity categories that were based on the Cambridge Taxonomy of Business Risks (CTBR). They defined values for BS, LR, epochs, MSL, and DR, but LR and BS were the only hyperparameters that were optimized in terms of their effect to the model's predictive accuracy and computational efficiency. LR and BS were found to have varying effects on different kinds of EOLLMs, indicating that a tailored approach is needed if EOLLMs are used for named entity recognition purposes.

Shishehgarkhaneh et al. (2024b) developed this implementation further and introduced a genetic algorithm (GA) for the hyperparameter tuning process for BERT and DeBERTa EOLLMs. Based on the SCRM context, they identified precision as a key performance metric to assess the model's performance in named entity recognition. Identified entities must as relevant as possible to the SC but minimizing the number of missed entities is not as key. Since the DeBERTa EOLLM achieved a greater precision score out of the two, the GA was used to tune the same hyperparameters as Shishehgarkhaneh et al. (2024a) did.

Although DOLLMs can be tailored to perform certain tasks with PE, better performance results can be achieved by FT them, especially in NLU tasks. FT of DOLLMs is however more complex than EOLLMs, as changing the parameters weights might lead to the DOLLM losing knowledge attained during the pre-training process and optimization of the large amount of DOLLM parameters is computationally very expensive. To address these issues, parameter-efficient FT of DOLLMs only optimizes few parameters that are relevant to the downstream task. (Feroze et al. 2025; Huang et al. 2025.) The same optimization hyperparameters are used in the FT of DOLLMs as in EOLLMs but there are also some additions. The number of **warmup steps** (WS) refers to the gradual increase in the LR when starting the fine-tuning process. The training process is started with a low value to avoid instabilities caused by a greater LR. (Huang et al. 2025.)

Sun et al. (2024) utilized a GPT-3.5 DOLLM and fine-tuned it to extract and classify SCRs and predict SCR trends from NAs. The DOLLM was optimized over LR, BS, epochs, and WSs. By utilizing FT, the model achieved better results in tasks related to the extraction of SCRs.

4.2 Utilizing LLMs to extract SCRs

4.2.1 Schema learning

Schema learning (SL) is based on the schema theory, a multidisciplinary research subject that aims to define the process of knowledge creation in humans and more recently in computers. A schema is based on relationships between components that represent different identities and events and represents the existing general knowledge of a given concept. SL is the process of linking a learning task to an existing schema and updating the schema. By utilizing the concept of SL, LLMs can be used create graphs that depict the relationships between different entities and events. (Bein & Niv 2025; Cheng et al. 2024; Xia et al. 2024.)

Cheng et al. (2024) utilized various DOLLMs to create a schema library of components in electric vehicle (EV) SCs. DOLLMs were used to create a schema graph of events contributing to SCRs that encompass the complex interdependencies of EVSCs. The schema was created based on a dataset that was compiled with the help of SCM experts that contained detailed information of different components of the EVSC. The possibility of human-curation of the schemas created by the DOLLMs was retained through a user interface, where users can examine, validate, and modify the created schema.

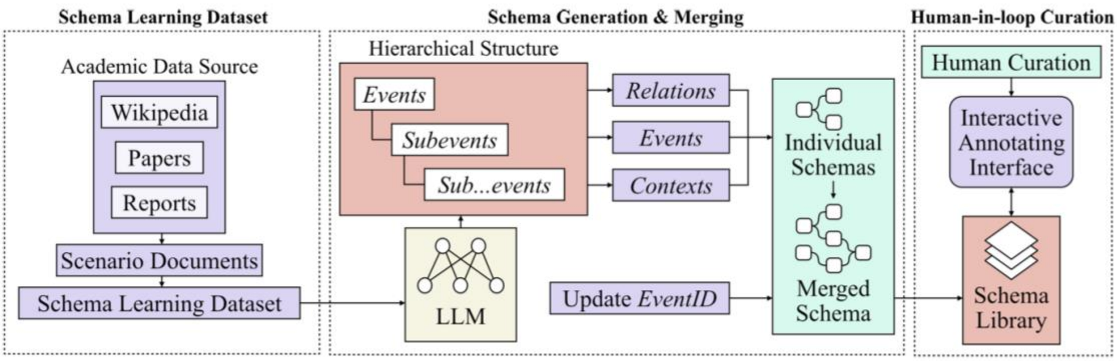

Figure 4. Schema construction process for events contributing to SCRs, figure copied from (Cheng et al. 2024)

4.2.2 Semantic text similarity analysis

Semantic text similarity analysis (STSA) is an NLP task that aims to define the semantic similarities between different text sources. DOLLMs have been found to outperform EOLLMs in STSA, but EOLLMs are still viable options for STSA if computational resources are limited. LLM-based STSA is implemented by giving the LLM text inputs and asking it to analyze the semantic differences between the texts. The semantic similarity is often calculated by for example the cosine similarities (CS) of the WEVs of the different text inputs. The CS is defined as the “cosine of the angle between two vectors” (Xia et al. 2015, 42) and can be stated with the formula $\cos(\theta) = \frac{A \cdot B}{\|A\| \times \|B\|}$, where $\cos(\theta)$ is the cosine angle and A and B are the two WEVs. The CS values range between -1 and 1, with a CS value of -1 indicating that the two vectors are exactly the opposite from each other and a CS value of 1 indicating that vectors are exactly the same. PE techniques, such as chain-of-thought prompting, are often used to achieve better STSA performance. (Benayas et al. 2024; Feng 2024; Hussain et al. 2023; Jin & Lin 2024.)

Cheng et al. (2024) calculated a composite similarity for extracted SCR events and the schemas introduced in section 4.2.1 that was based on the semantic and structural similarities of the texts. Semantic similarity was calculated based on the CS of the texts and structural similarity was calculated based on the Jaccard similarity of the texts. The event was categorized based on the schema event category with the highest similarity score with the event.

Shahsavari et al. (2024a) and Shahsavari et al. (2024b) utilized a pre-trained sentence transformer EOLLM in their framework to compute CSs between the name of an event contributing to SCRs and various texts. A more detailed description of how these LLMs were utilized is given in section 4.3.

Zhao, M. et al. (2024a) utilized open-source pre-trained EOLLMs to classify the SCRs that were identified with the GPT-3.5 DOLLM. The identified SCR text output and the 875 business risk definitions from the CTBR were given to EOLLMs as inputs and the EOLLMs were tasked to identify the business risk that was most similar with the identified SCR, based on the CSs of the texts’ WEVs. The more complex EOLLMs, such as sentence-t5-base, achieved better labeling accuracy than the simpler EOLLMs and were able to classify the correct label even when the generated description of the SCR was non-explicit.

4.2.3 Key phrase extraction

Key phrase extraction (KPE) is an NLP task that aims to extract key phrases that describe and express the central content of textual documents. KPE has many use cases, for example in generating search queries, identifying recurring phrases in academic research or transforming key phrases to vector-representation for document classification and clustering. Both EOLLMs and DOLLMs can be utilized in KPE. (Kang & Shin 2025; Yao et al. 2022.)

Shahsavari et al. (2024a) and Shahsavari et al. (2024b) utilized the KeyBERT EOLLM to extract key phrases from NAs in their framework.

4.2.4 Event extraction

Event extraction (EE) is an NLP task that aims to detect events from text sources. EE is performed through semantic analysis of text sources to detect triggers for event, entities that are related to the event and other information related to the event, such as time and place. (Meng et al. 2024.)

Cheng et al. (2024) utilized a fine-tuned RoBERTa EOLLM to detect distribution events relating to EVSCs from NAs. Events were extracted based on sequence tagging to conduct cross-sentence semantic analysis of the NAs.

Shahsavari et al. (2024a) and Shahsavari et al. (2024b) utilized a GPT-3.5 DOLLM in their framework for EE.

4.2.5 Named entity recognition

Named entity recognition (NER) is an NLP task that aims to identify predefined entity types from text documents. NER is a key task in most IE systems and aids in downstream tasks such as relation extraction. LLMs have shown their considerable potential to be utilized in NER. (Jung et al. 2024; Shishehgarkhaneh et al. 2024a)

Shishehgarkhaneh et al. (2024a) and Shishehgarkhaneh et al. (2024b) utilized EOLLMs for NER to extract SCRs from NAs.

Sun et al. (2024) utilized NER in pre-processing of FT training data to standardize the mentions of different entities.

4.3 Integrating LLMs extracting SCRs to the SCRM process

The conventional SCRM process has three main issues in the modern SCE: it’s too reactive in nature to be used in the modern SCE, that requires proactive risk management processes, it relies too heavily on the availability of accurate historical data, and its need for domain-specific expertise, different industries have industry-specific SCRs, meaning that all general SCRM processes require industry-specific adaptations. (Aljohani 2023; Deiva Ganesh & Kalpana 2022; Kırılmaz & Erol 2017; Mukherjee et al. 2024; Srivastava & Rogers 2022;

Zhao, M. et al. 2024a; Zhao, M. et al. 2024b.) Integrating LLMs to the SCRM process can be used to solve these issues.

Cheng et al. (2024) introduced an implementation that utilized various LLMs and other AI technologies to extract and assess the impact of disruption events that contribute to SCRs in EVSCs. The schema of SCRs in EVSC was first created, as introduced in section 4.2.1. Consequently, a dataset of NAs about SCR events in EVSCs was created using a Python web crawler. The dataset was cleaned by using a GPT-4o DOLLM and manual curation to assess the NAs relevance to the EVSC and a Llama3-8b to filter out unnecessary information, such embedded advertisements. In the implementation pipeline, events are first extracted from this dataset, as introduced in section 4.2.4. After EE, a BERT EOLLM is used to generate contextual embeddings to enrich the contextual information of the extracted event. Consequently, a RoBERTa EOLLM fine-tuned with the TempEval3 dataset is used to extract time expressions related to the event to define when the event has happened. By using a fine-tuned RoBERTa-large EOLLM, extracted events are set in chronological order to enable the next step. Coreference resolution (CR) and event linking between extracted events mentioned in single and multiple documents is used to create a logical event sequence of events contributing to SCRs. The linked events are analyzed using conditional random fields (CRFs) to extract parameters contributing to the events. The impact of each event is assessed based on the event's magnitude and its centrality to the EVSC by using a graph convolutional network (GCN). To ensure the robustness of created event sequences, logical constraints (LC) and CR are introduced for event sequences. The constraints define that for example events with chronological dependencies require that event A is recognized to happen before event B or that event A is recognized as a precursor of event B. CR ensures the within- and cross-document coreference consistency of event mentions. By combining LCs and CR, the pipeline creates high-quality data about events contributing to SCRs. The extracted events are then categorized, as introduced in section 4.2.2. Ensuring that the categorization adheres to the logical and temporal constraints introduced earlier improves its accuracy. To ensure continuous improvement of the pipeline, industry experts assess the quality of the extracted events and give feedback that is used to refine and update the model. The extraction of events also allows the pipeline to be used to predict future disruption events. Predicted event sequences are based on the propagation of events in the created event schema. Future disruption events are predicted by utilizing a GCN trained to learn schema graph nodes and edges, by enforcing LCs on, for example, the propagation of child and parent nodes, and ensuring consistent entity mentions through argument coreference. Overall, the pipeline can be used through an online interface that allows the user to input their own NAs for EE, assess the quality of extracted events, and modify event sequence visualizations created by the pipeline.

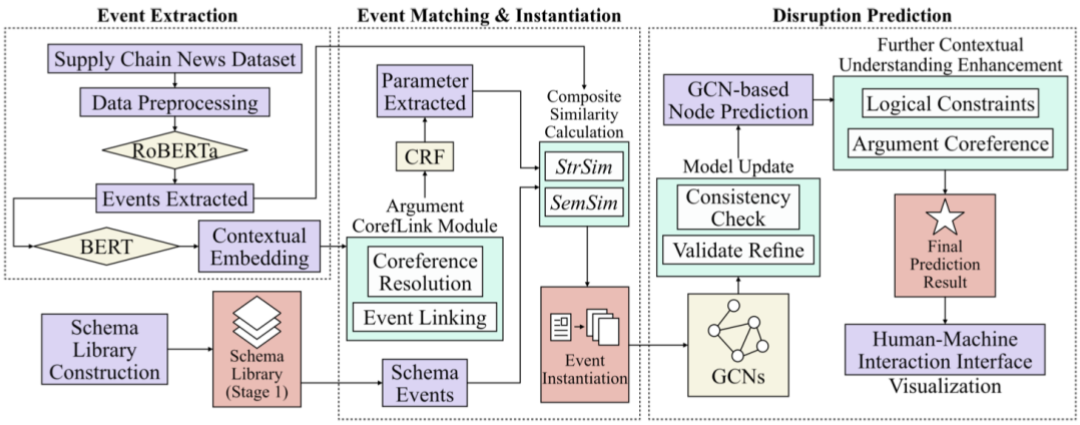

Figure 5. Pipeline of implementation, figure copied from (Cheng et al. 2024)

Fan et al. (2025) introduced an implementation that utilized a GPT-4o mini DOLLM to measure the company's exposure to SCRs from the transcriptions of corporate site visit audits. Investors and company shareholders can initiate these audits to assess to current business situation of the company. PE was utilized in this implementation, as introduced in section 4.1.1. The company's exposure to the identified SCR was assessed based on the length of the audit transcription and the length of the response generated by the DOLLM.

Shahsavari et al. (2024a) and Shahsavari et al. (2024b) introduced a SCR EE framework consisting of six modules for identifying events that contribute to SCRs. The implementation utilized LLMs for STSA and IE purposes. The framework first creates causal relationships between SCRs and events that contribute to them by analyzing historical news data, identifying causal relationships using Bayesian networks, and creating key phrases that describe the CE and their relationship with the SCR. WordNet, a lexical database of English words and their synonyms, is used in the next module to find synonyms for single words of the key phrases (e.g. staff and stave are synonyms and shortage and shortfall are synonyms), to create synonym combinations of the key phrases (e.g. staff shortage and stave shortage). Some of the provided synonym combinations, however, do not have the same semantic meaning (e.g. staff shortage and stave shortage). SBERT is used to transform the sentences into WEVs and calculate the CS of the given sentences and the name of the CE. If the CS value is 0.8 or greater, the synonym combination is accepted as a search phrase. Further search phrases are searched for by first searching for NAs published within the last three years and then using the accepted synonym combinations to search for the articles. If new NAs are found using the synonym combinations, it is accepted as a final search phrase but if no new articles are found, the synonym combination is dropped. The extracted NAs are analyzed using KeyBERT, as described in section 4.2.3. CSs are calculated using SBERT for the key phrases and the name of the CE. If the CS value is 0.8 or greater, the extracted phrase is accepted as a final search phrase. The final search phrases are input to the next module, a news crawler that utilizes Python's pygooglenews API to search for NAs from Google News. Based on the search results, CSs are calculated using SBERT for the CE and the extracted NAs. If the calculated CS value is 0.5 or greater, the NA is taken into further consideration, otherwise it is dropped as irrelevant to the CE. The CS values between a CE and NAs contribute to the probability of occurrence of the CE, a CS value in the range of 0.5 and 0.57 is considered low probability, a value in the range of 0.57 and 0.64 is considered medium probability, and a value greater than 0.64 is considered high probability. The relative frequencies of the different probabilities are used as the probability of the CE having a certain probability (e.g. 1 out of 10 NAs have a CS value between 0.5 and 0.57, the CE is deemed to have a 10% probability of having a low probability of occurence). The relevant NAs are analyzed to extract key information about the CE, as introduced in section 4.1.1. The final module propagates the calculated probabilities of CEs through the Bayesian network generated in the first module to generate an estimate on the probability of occurrence for the SCR in question.

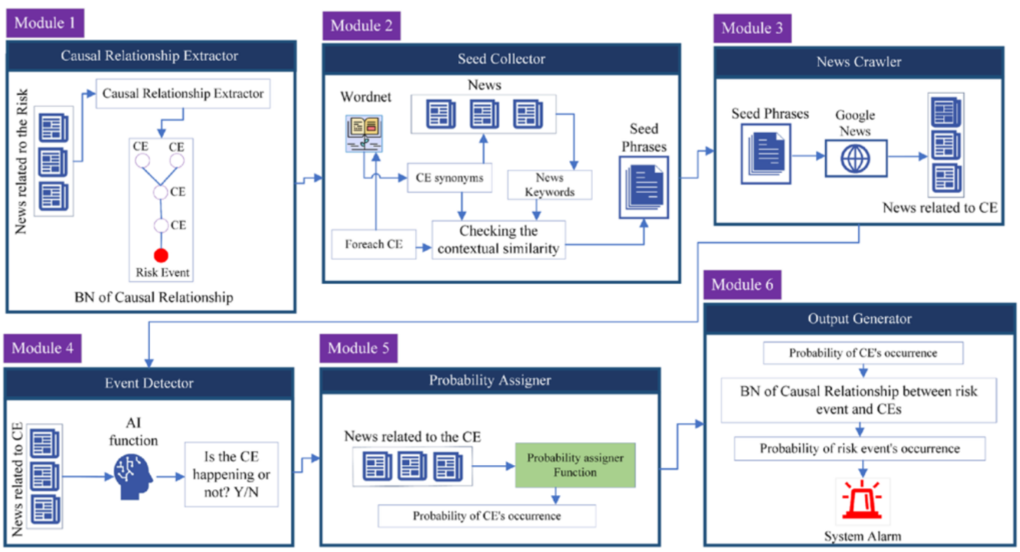

Figure 6. Contributing Event-based Risk Identification and Assessment -framework, figure copied from (Shahsavari et al. 2024a; Shahsavari et al. 2024b)

Shishehgarkhaneh et al. (2024a) and Shishehgarkhaneh et al. (2024b) introduced an implementation that utilized fine-tuned EOLLMs for NER purposes to extract SCRs from NAs. Based on the FT introduced in section 4.1.2, the DeBERTa EOLLM was identified as the optimal EOLLM for this use case.

Sun et al. (2024) introduced an implementation based on a fine-tuned GPT-3.5 DOLLM that extracted SCRs and used the extracted SCRs to predict future SCR trends. The DOLLM was fine-tuned as described section 4.1.2. The fine-tuned DOLLM was used to extract SCR categories from NAs and then extract and classify SCRs from NAs identified SCR categories. Using the date information of the extracted SCRs, the DOLLM was used to predict the future timeline of the extracted SCR and assess the impact it would have on the SC in the future.

Zhao, M. et al. (2024a) and Zhao, M. et al. (2024b) introduced a Python-based prototype software that could be used to support and solve the issues with the conventional SCRM process. The software is setup by inserting the geographical locations (e.g. country, region) of the focal company and all known suppliers in the focal company's SC. The input information is used in all the concurrent steps. The system utilizes a web crawler that automatically retrieves real-time news that discuss their suppliers or their geographical locations from news databases, such as Google News, allowing for a more proactive identification of SCRs that does not require accurate historical data to function. These NAs consequently analyzed, as introduced in sections 4.1.1 and 4.2.2. The extracted SCRs are visualized using Neo4j, a graph database system. By utilizing Neo4j, the links between the SCRs, suppliers, and the focal company are clearly visualized, so that SCRM decision-makers can easily review what kind of impact the SCRs could possibly have on the SC and what is the possibility of them happening.

Based on the above-mentioned implementations, it can be concluded that LLMs can be used to support all stages of the SCRM process. By utilizing LLMs, SCRs can be identified, their effects can be assessed, and they can be monitored for shifts in trends. LLMs can also create invaluable information that can be used in the creation of SCR mitigation strategies. NAs contain relevant information that can be used to monitor individual SCRs, but also the broader shifts in the SCE (Chu et al. 2020). These implementations allow for a robust SCRM process in uncertain SCEs, where accurate data is scarce, by transforming unstructured textual data into structured data that can be utilized in the SCRM process. The introduced implementations also show that LLMs can be used for robust proactive SCRM better than conventional SCRM methods.

4.4 Future development of LLMs in identifying SCRs

Despite that the implementations introduced in section 4.3 allow for a more proactive SCRM process, they can still be improved. Most of the implementations introduced tasked the LLMs in use to label the extracted SCRs based on the CTBR definitions. This also true for other SCR extraction implementations that utilize other AI technologies than LLMs (Aboutorab et al. 2022; Aboutorab et al. 2023; Aboutorab et al. 2024). Only Shishehgarkhaneh et al. (2024a) and Shishehgarkhaneh et al. (2024b) and Cheng et al. (2024) defined a taxonomy of SCRs themselves, but the taxonomies they created were tailored for SCs in specific industries, construction and EVs respectively. The need for industry-tailored taxonomies shows that many SCRs can be industry-specific.

Searching Google Scholar with the search prompt “supply chain” AND “Cambridge taxonomy” only returns 17 articles, with most articles utilizing the CTBR definitions in AI implementations and no articles on conventional SCRM themes. When given the prompt “relationship of supply chain risk management and Cambridge taxonomy of business risks”, SCOPUS AI and Consensus AI search engines also returned articles with similar results. The results indicate that there exists a gap between the AI-based SCRM literature and the conventional SCRM literature. Because there are no mentions of utilizing the CTBR in conventional SCRM literature, it indicates that the CTBR definitions are not detailed enough so that could be directly utilized in SCRM. The CTBR has six main branches of risks: financial, geopolitical, technology, environmental, social and governance (Cambridge Centre for Risk Studies 2019) while Table 1 already has 12 defined SCR types, with numerous other definitions found in SCRM literature. But because the AI implementations require organized definitions of SCRs that have quantifiable impact, AI-researchers have opted for the CTBR definitions because of the difficulty to define and lack of an organized taxonomy of SCRs (Fan et al. 2025; Pandey et al. 2023; Sun et al. 2023). For example, Fan et al. (2025) only used three SCR factors to determine a SC’s exposure to SCRs.

Based on this, it is clear that a more general organized taxonomy of SCRs, similar to the ones defined by Shishehgarkhaneh et al. (2024a) and Cheng et al. (2024), is needed to develop the capabilities of AI implementations in identifying SCRs. The CTBR can be used to draw insights, such as Shishehgarkhaneh et al. (2024a) and Shishehgarkhaneh et al. (2024b) did, but using the CTBR definitions solely does not accurately reflect the modern SCRs. It is important that SCM experts are consulted in the creation of this taxonomy, similar to Cheng et al. (2024) who created their SCR categorization with the help of industry experts.

Due to the need for the inclusion of industry-specific SCRs (Srivastava & Rogers 2022), the general taxonomy could be used as a basis for an industry-specific LLM-based implementation. The SCRs that are relevant to the given industry could be picked from the general taxonomy, while non-relevant SCRs would be left out. General taxonomy SCRs could also be tailored to include industry-specific characteristics.

The utilized data sources should also be diversified to better reflect the modern SCRs. Although NAs contains valuable insights on SCRs, they themselves cannot capture all the complex interdependencies in SCs. The implementation introduced by Fan et al. (2025) that utilized a private audit transcript was the only implementation that utilized an internal data source. Internal data sources, especially numerical metrics such as supply chain performance metrics, should be better integrated to the SCR extraction process in the future (Sun et al. 2024). The variety of public data sources should also be diversified. Social media posts, industry reports and government publications are some examples of data sources that could be utilized to extract SCRs (Shishehgarkhaneh et al. 2024b).

5 Discussion and summary

5.1 Discussion

Sections 4.1, 4.2 and 4.3 provide an overview of how LLMs can be used to detect SCRs. In response to RQ1, it can be concluded that LLMs are capable of detecting SCRs through the application of diverse techniques across multiple NLP tasks. For DOLLMs, PE and hyperparameter tuning are the main techniques used to tailor them for SCR identification and for EOLLMs, FT is the most prevalent technique. STSA, KPE, EE, NER and coreference are some of the NLP tasks that can be used to create an implementation for SCR identification. LLMs can also be seamlessly integrated with other AI technologies in implementations, such as CRFs and GCNs that Cheng et al. (2024) utilized in their implementation. Most implementations are based on identifying SCR events from external textual data sources, such as NAs, meaning that these kinds of implementations would be easiest for companies to implement themselves.

One interesting result based on the articles introduced in Table 2 is that almost none of them specify the types of SCRs that are reliably extractable. Only Cheng et al. (2024) specify the SCR event categories that are extracted using their implementation. Political, demand, and operational SCRs are mentioned examples of extracted SCRs in the article, but the classification accuracy for every category is not elaborated. This means in response to RQ2, it can be concluded that the lack of specification for the classification accuracies of single SCRs makes it difficult define SCRs that are the most reliably extractable using SCRs.

In response to RQ3, it can be concluded that based on the implementations introduced in section 4.3, LLMs can be used to support all stages of the SCRM process, meaning that LLM technologies have considerable use cases in the field of SCRM. The main benefit of utilizing these methods is their robustness in uncertain SCEs. All implementations include a measure for the quality of extracted information, ensuring that implementations are capable of filtering out inaccurate data. Some implementations also offer a possibility for visualization for the extracted SCRs, improving the visibility of the supply chain, which is a big problem in modern SCs (Kassa et al. 2023). The implementation introduced by Zhao et al. (2024) features a visualization of the complex interdependencies that contribute to SCRs, helping decision-makers grasp them better. The implementations introduced by Cheng et al. (2024) and Shahsavari et al. (2024a) and Shahsavari et al. (2024b) include event sequences that contribute to SCRs, aiding in the creation of mitigation strategies in SCRM frameworks that are based on the analysis of events contributing to SCRs, such as the HOR method. The introduced implementations offer a proactive and easy-to-use option with their user interfaces to support the conventional SCRM methods in the face of modern dynamic SCEs.

However, between LLM and SCRM literature there exists a gap, hinted by the lack of cross-domain co-operation. Only Cheng et al. (2024) and Shahsavari et al. (2024a) and Shahsavari et al. (2024b) co-operated with SCRM domain experts in their implementations, hinting that that more can be achieved with LLMs in a SCRM with the co-operation of different domain experts. A general taxonomy would clearly differentiate the characteristics of different SCRs, enabling the examination of the classification accuracies of LLMs for different SCRs. The uneven prevalence of different attributes can lead to the LLM having varying identification performance for different entities (Keloth et al. 2024). Each SCR category has its own attributes and characteristics, meaning that it is plausible that LLMs could have better knowledge on the characteristics of certain SCRs and could achieve better performance in identifying those SCRs that they have better knowledge on. In response to RQ4, it can be concluded that there is a clear need for a generalized taxonomy of SCRs that can be used with LLMs, created through co-operation between SCRM and LLM domain experts. In industry-specific implementations, the general taxonomy could be used as a basis for the used SCR categorization, as relevant SCRs could be picked from the general taxonomy and tailored to include industry-specific characteristics if needed. Creating the taxonomy would also help in identifying the SCRs that are most reliably extractable using LLMs. The focus in future research should also be in diversifying the utilized data sources. The utilization of internal data sources should be more prevalent in future implementations, as they contain a considerable amount data about SCs.

5.2 Summary

SCs are very large networks spanning global areas and contain complex interdependencies between different actors. As the complexity of SCs is ever-increasing, SCs are becoming vulnerable towards an increasing number of SCRs. SCRM is the process of managing these risks, revolving around identifying SCRs, assessing their impact, creating mitigation strategies for them and continually monitoring the risks and the effect of created mitigation strategies. Conventional SCRM methods however have a hard time keeping up with the large number of SCRs, contributing to a need for more responsive SCRM methods. Due to complex nature of SCs, SCRs can be categorized in many different ways, making it difficult to create a generally applicable categorization of SCRs for SCRM.

LLMs are state-of-art AI models that are subject to growing academic interest. LLMs have unrivaled NLU and NLG capabilities that help them to perform various unformalized tasks. Thanks to their vast training dataset, LLMs have acquired very specific domain knowledge that can be harnessed with the right tools, such as PE and FT. By utilizing these tools, LLMs have been successfully trialed in many academic fields, including SCRM.

By tailoring LLMs for SCRM purposes, they can be used to support the conventional SCRM process and make it more proactive and robust in modern uncertain SCEs. In the future, AI and SCRM domain experts should co-operate in creating a generalized taxonomy of SCRs to allow for more precise evaluations of their identification capabilities and to improve the capabilities of LLMs in the SCRM context.

References

Aboutorab, Hamed; Hussain, Omar K.; Saberi, Morteza; Hussain, Farookh Khadeer (2022) A reinforcement learning-based framework for disruption risk identification in supply chains. Future Generation Computer Systems 126, 110–122

Aboutorab, Hamed; Saberi, Morteza; Hussain, Omar K.; Hussain, Farookh Khadeer (2023) POSSUM: PrOactive diSruption riSk identification for sUpply chain Management. IEEE International Conference on Pervasive Computing and Communications Workshops and other Affiliated Events (PerCom Workshops), 319–321

Aboutorab, Hamed; Hussain, Omar K.; Saberi, Morteza; Hussain, Farookh Khadeer; Prior, Daniel (2024) Adaptive identification of supply chain disruptions through reinforcement learning. Expert Systems With Applications 248, 123477

Aghaei, Raha; Kiaei, Ali A; Boush, Mahnaz; Vahidi, Javad; Barzegar, Zeynab; Rofoosheh, Mahan (2025) The Potential of Large Language Models in Supply Chain Management: Advancing Decision-Making, Efficiency, and Innovation. arXiv preprint arXiv:2501.15411

Ahammad, Tanvir; Al Mamun, Md Sajib; Tabassum, Mehnaz (2016) Towards the Application of Big Data: A New Way to make Data Driven Healthcare Decision. International Journal of Computer Applications 134(14), 15–21

Aljohani, Abeer (2023) Predictive Analytics and Machine Learning for Real-Time Supply Chain Risk Mitigation and Agility. Sustainability 15(20), 15088

Almatarneh, Zeyad; Jarah, Baker Akram Falah; Jarrah, Mufleh Amin AL (2022) The role of management accounting in the development of supply chain performance in logistics manufacturing companies. Uncertain Supply Chain Management 10(1), 13–18

Ang, Phyllis; Dhingra, Bhuwan; Wills, Lisa Wu (2022) Characterizing the Efficiency vs. Accuracy Trade-off for Long-Context NLP Models. Proceedings of NLP Power! The First Workshop on Efficient Benchmarking in NLP, 113–121

Badwan, Nemer (2024) Role of supply chain partnership, cross-functional integration, responsiveness and resilience on competitive advantages: empirical evidence from Palestine. TQM Journal, 2024-05

Baryannis, George; Validi, Sahar; Dani, Samir; Antoniou, Grigoris (2018) Supply chain risk management and artificial intelligence: state of the art and future research directions. International Journal of Production Research 57(7), 2179–2202

Beheshti-Kashi, Samaneh; Pannek, Jürgen; Kinra, Aseem (2019) Complementing Decision Support and Forecasting Risk in Supply Chain with Unstructured Data. IFAC-PapersOnLine 52(13), 1721–1726

Bein, Oded; Niv, Yael (2025) Schemas, reinforcement learning and the medial prefrontal cortex. Nature Reviews Neuroscience 26(3), 141–157

Bellomi, Francesco; Cristani, Matteo (2024) Text classification for private procurement: a survey and an analysis of future trends. Information Technology and Management, 2024-11

Benayas, Alberto; Sicilia, Miguel Angel; Mora-Cantallops, Marçal (2024) A comparative analysis of encoder only and decoder only models in intent classification and sentiment analysis: navigating the trade-offs in model size and performance. Language Resources and Evaluation.

Blos, Mauricio F.; Quaddus, Mohammed; Wee, H.M.; Watanabe, Kenji (2009) Supply chain risk management (SCRM): a case study on the automotive and electronic industries in Brazil. Supply Chain Management 14(4), 247–252

Bodendorf, Freimut; Zimmermann, Roland (2005) Proactive Supply-Chain Event Management with Agent Technology. International Journal of Electronic Commerce 9(4), 58–89

Caillaut, Gaëtan; Qader, Raheel; Nakhlé, Mariam; Liu, Jingshu; Barthélemy, Jean-Gabriel (2024) Scaling Laws of Decoder-Only Models on the Multilingual Machine Translation Task. Proceedings of the Ninth Conference on Machine Translation, 1318–1331

Cambridge Centre for Risk Studies (2019) Cambridge Taxonomy of Business Risks

Chang, Hsihui; Ittner, Christopher D.; Paz, Michael T. (2014) The Multiple Roles of the Finance Organization: Determinants, Effectiveness, and the Moderating Influence of Information System Integration. Journal of Management Accounting Research 26(2), 1–32

Cheng, Zhi-Qi; Dong, Yifei; Shi, Aike; Liu, Wei; Hu, Yuzhi; O'Connor, Jason; Hauptmann, Alexander G; Whitefoot, Kate S (2024) SHIELD: LLM-Driven Schema Induction for

Predictive Analytics in EV Battery Supply Chain Disruptions. Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing: Industry Track, 303–333

Chu, Chih-Yuan; Park, Kijung; Kremer, Gül E. (2020) A global supply chain risk management framework: An application of text-mining to identify region-specific supply chain risks. Advanced Engineering Informatics 45, 101053

Dang, Yizhou; Liu, Yuting; Yang, Enneng; Guo, Guibing; Jiang, Linying; Wang, Xingwei; Zhao, Jianzhe; Verbert, Katrien; Lops, Pasquale; Noia, Tommaso Di; Joachims, Thorsten; Dong, Zhenhua; Castells, Pablo; London, Ben (2024) Repeated Padding for Sequential Recommendation. Proceedings of the 18th ACM Conference on Recommender Systems, 497–506

Davis, Joshua; Van Bulck, Liesbet; Durieux, Brigitte N; Lindvall, Charlotta (2024) The Temperature Feature of ChatGPT: Modifying Creativity for Clinical Research. JMIR Human Factors 11, e53559

de Oliveira, Ualison Rébula; Marins, Fernando Augusto Silva; Rocha, Henrique Martins; Salomon, Valério Antonio Pamplona (2017) The ISO 31000 standard in supply chain risk management. Journal of Cleaner Production 151, 616–633

Deiva Ganesh, A.; Kalpana, P. (2022) Future of artificial intelligence and its influence on supply chain risk management; A systematic review. Computers & Industrial Engineering 169, 108206

Delgado-Verde, Miriam; Díez-Vial, Isabel (2024) New product development and supplier involvement: the role of R&D collaboration with supporting organisations. Journal of Technology Transfer 49(2), 518–541

Dewett, Todd; Jones, Gareth R. (2001) The role of information technology in the organization: a review, model, and assessment. Journal of Management 27(3), 313–346

Ding, Tianyi; Huang, Zongsheng (2024) Uncovering the Research Hotspots in Supply Chain Risk Management from 2004 to 2023: A Bibliometric Analysis. Sustainability 16(12), 526

Dukić, David; Šnajder, Jan (2024) Looking Right is Sometimes Right: Investigating the Capabilities of Decoder-only LLMs for Sequence Labeling. Findings of the Association for Computational Linguistics: ACL 2024, 14168–14181

Fan, Siyu; Wu, Yifei, Wu; Yang, Ruochen (2025) Measuring firm-level supply chain risk using a generative large language model. Finance Research Letters 77, 107111

Feng, Yeli (2024) Semantic Textual Similarity Analysis of Clinical Text in the Era of LLM. 2024 IEEE Conference on Artificial Intelligence, 1284–1289

Fiksel, Joseph; Polyviou, Mikaella; Croxton, Keely L.; Pettit, Timothy J. (2015) From Risk to Resilience: Learning to Deal With Disruption. MIT Sloan Management Review 56(2), 78–86

Feroze, Wasif; Cheng, Shaohuan; Jimale, Elias Lemuye; Jakhro, Abdul Naveed; Qu, Hong (2025) Enhancing text understanding of decoder-based model by leveraging parameter-efficient fine-tuning method. Neural Computing & Applications 37(9), 6899–6913

Forbes, Naushad; Wield, David (2004) What is R&D? Why does it matter?. Science & Public Policy 31(4), 267–277

Gezdur, Arda; Bhattacharjya, Jyotirmoyee (2025) Innovators and transformers: enhancing supply chain employee training with an innovative application of a large language model. International Journal of Physical Distribution & Logistics Management, 2025-01

Gómez-Cedeño, Milena; Castán-Farrero, José María; Guitart-Tarrés, Laura; Matute-Vallejo, Jorge (2015) Impact of human resources on supply chain management and performance. Industrial Management & Data Systems 115(1), 129–157

Grandi, Daniele; Patawari Jain, Yash; Groom, Allin; Cramer, Brandon; McComb, Christopher (2025) Evaluating Large Language Models for Material Selection. Journal of Computing and Information Science in Engineering 25(2), 021004

Guenzi, Paolo; Troilo, Gabriele (2007) The joint contribution of marketing and sales to the creation of superior customer value. Journal of Business Research 60(2), 98–107

Haq, Muhammad Zia Ul; Gu, Minhao; Huo, Baofeng (2021) Enhancing supply chain learning and innovation performance through human resource management. Journal of Business and Industrial Marketing 36(3), 552–568

Helmold, Mark; Samara, Warda (2019) Progress in performance management. Springer International Publishing

Hermoso-Orzáez, M. J.; Garzón-Moreno, J. (2021) Risk management methodology in the supply chain: a case study applied. Annals of Operations Research 313(2), 1051–1075

Ho, William; Zheng, Tian; Yildiz, Hakan; Talluri, Srinivas (2015) Supply chain risk management: a literature review. International Journal of Production Research 53(16), 5031–5069

Huang, Zixian; Huang, Xinwei; Wu, Ao; Wang, Xiaxia; Cheng, Gong (2025) Transforming decoder-only models into encoder-only models with improved understanding capabilities. Knowledge-Based Systems 309, 112907

Hussain, Musarrat; Rehman, Ubaid Ur; Nguyen, Tri D.T.; Lee, Sungyoung (2023) CoT-STS: A Zero Shot Chain-of-Thought Prompting for Semantic Textual Similarity. Proceedings of the 2023 6th Artificial Intelligence and Cloud Computing Conference, 135–139

Ito, Tomoki; Nakagawa, Shun (2024) Tender Document Analyzer with the Combination of Supervised Learning and LLM-based Improver. Companion Proceedings of the ACM Web Conference 2024, 995–998

Jin, Xin; Lin, Zhiqiang (2024) SimLLM: Calculating Semantic Similarity in Code Summaries using a Large Language Model-Based Approach. Proceedings of the ACM on Software Engineering 1(FSE), 1376–1399

Jung, Sung Jae; Kim, Hajung; Jang, Kyoung Sang (2024) LLM based Biological Named Entity Recognition from Scientific Literature. 2024 IEEE International Conference on Big Data and Smart Computing, 433–435

Jüttner, Uta; Peck, Helen; Christopher, Martin (2003) Supply chain risk management: outlining an agenda for future research. International Journal of Logistics 6(4), 197–210

Kang, Byungha; Shin, Youhyun (2025) Empirical Study of Zero-shot Keyphrase Extraction with Large Language Models. Proceedings of the 31st International Conference on Computational Linguistics, 3670–3686

Kassa, Adane; Kitaw, Daniel; Stache, Ulrich; Beshah, Birhanu; Degefu, Getachew (2023) Artificial intelligence techniques for enhancing supply chain resilience: A systematic literature review, holistic framework, and future research. Computers & Industrial Engineering 186, 109714

Keloth, Vipina K; Hu, Yan; Xie, Qianqian; Peng, Xueqing; Wang, Yan; Zheng, Andrew; Selek, Melih; Raja, Kalpana; Wei, Chih Hsuan; Jin, Qiao; Lu, Zhiyong; Chen, Qingyu; Xu, Hua (2024) Advancing entity recognition in biomedicine via instruction tuning of large language models. Bioinformatics 40(4), btae163

Khan, Syed Abdul Rehman; Yu, Zhang (2019) Strategic supply chain management. Springer, Switzerland

Kırılmaz, Oguzhan; Erol, Serpil (2017) A proactive approach to supply chain risk management: Shifting orders among suppliers to mitigate the supply side risks. Journal of Purchasing & Supply Management 23(1), 54–65

Kitchin, Rob; McArdle, Gavin (2016) What makes Big Data, Big Data? Exploring the ontological characteristics of 26 datasets. Big Data & Society 3(1)

Kleindorfer, Paul R.; Saad, Germaine H. (2005) Managing Disruption Risks in Supply Chains. Production and Operation Management 14(1), 53–68

Kühl, Linus; Wiethöler, Jost; Dircksen, Michael (2024) Enhancing Supply Chain Risk Identification: Analyzing the Impact of LLM Parameters for precise Classification. Proceedings of the 28th International Symposium on Logistics, 197–205

Lambert, Douglas M. (2008) Supply chain management: processes, partnerships, performance. 3rd Edition. Suppy Chain Management Institute, Sarasota, Florida

Li, Zhihong; Ksibi, Albaraa; Xu, Xiaoying (2024) Optimizing Inventory Management using a Multi-Agent LLM System. The 24th International Conference on Electronic Business, 308–318

Livne, Micha; Miftahutdinov, Zulfat; Tutubalina, Elena; Kuznetsov, Maksim; Polykovskiy, Daniil; Brundyn, Annika; Jhunjhunwala, Aastha; Costa, Anthony; Aliper, Alex; Aspuru-Guzik, Alán; Zhavoronkov, Alex (2024) nach0: multimodal natural and chemical languages foundation model. Chemical Science 15(22), 8380–8389

Meng, Zihao; Liu, Tao; Zhang, Heng; Feng, Kai; Zhao, Peng (2024) CEAN: Contrastive Event Aggregation Network with LLM-based Augmentation for Event Extraction. Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics Volume 1: Long Papers, 321–333

Mentzer, John T.; DeWitt, William; Keebler, James S.; Min, Soonhong; Nix, Nancy W.; Smith, Carlo D.; Zacharia, Zach G. (2001) Defining supply chain management. Journal of Business Logistics 22(2), 1–25

Mentzer, John T.; Stank, Theodore P.; Esper, Terry L. (2008) Supply Chain Management and its relationship to logistics, marketing, production and operations management. Journal of Business Logistics 29(1), 31–46

Mukherjee, Swarup; De, Anupam; Roy, Supriyo (2024) Supply chain risk prioritization: a multi-criteria based Intuitionistic Fuzzy TOPSIS approach. International Journal of Quality & Reliability Management 41(6), 1693–1725

Nakano, Mikihisa (2019) Supply Chain Management. 1st Edition. Springer, Singapore

Ni, Shiwen; Yang, Min; Xu, Ruifeng; Li, Chengming; Hu, Xiping (2024) Layer-wise Regularized Dropout for Neural Language Models. Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation, 10208–10218

Norrman, Andreas; Jansson, Ulf (2004) Ericsson's proactive supply chain risk management approach after a serious sub-supplier accident. International Journal of Physical Distribution & Logistics Management 34(5), 434–456

Nyoman Pujawan, I; Geraldin, Laudine H (2009) House of risk: a model for proactive supply chain risk management. Business Process Management Journal 15(6), 953–967

Olson, David L. (2014) Supply Chain Risk Management. 2nd Edition. Business Expert Press, New York

Omar, Mahmud; Brin, Dana; Glicksberg, Benjamin; Klang, Eyal (2024) Utilizing natural language processing and large language models in the diagnosis and prediction of infectious diseases: A systematic review. American Journal of Infection Control 52(9), 992–1001

Pandey, Shipra; Singh, Rajesh K.; Gunasekaran, Angappa (2023) Supply chain risks in Industry 4.0 environment: review and analysis framework. Production Planning & Control 34(13), 1275–1302

Pham, Hai Thanh; Pham, Tho; Truong Quang, Huy; Dang, Chau Ngoc (2023) Supply chain risk management research in construction: a systematic review. International Journal of Construction Management 23(11), 1945–1955

Polat, Fina; Tiddi, Ilaria; Groth, Paul (2025) Testing prompt engineering methods for knowledge extraction from text. Semantic Web 16(2), 1–32

Qiu, Zi-Hao; Guo, Siqi; Xu, Mao; Zhao, Tuo; Zhang, Lijun; Yang, Tianbao (2024) To Cool or not to Cool? Temperature Network Meets Large Foundation Models via DRO. Proceedings of the 41st International Conference on Machine Learning, 41604–41643

Raiaan, Mohaimenul Azam Khan; Mukta, Md. Saddam Hossain; Fatema, Kaniz; Fahad, Nur Mohammad; Sakib, Sadman; Mim, Most Marufatul Jannat; Ahmad, Jubaer; Ali, Mohammed Eunus; Azam, Sami (2024) A Review on Large Language Models: Architectures, Applications, Taxonomies, Open Issues and Challenges. IEEE Access 12, 26839–26874

Rangel, Djalma Araújo; de Oliveira, Taiane Kamel; Leite, Maria Silene Alexandre (2015) Supply chain risk classification: discussion and proposal. International Journal of Production Research 53(22), 6868–6887

Roosan, Don; Wu, Yanting; Chok, Jay; Sanine, Christopher P.; Khou, Tiffany; Li, Yawen; Khan, Hasiba M. (2024) Artificial Intelligence-Powered Large Language Transformer Models for Opioid Abuse and Social Determinants of Health Detection for the Underserved Population. Proceedings of the 13th International Conference on Data Science, Technology and Applications, 15–26

Rosoł, Maciej; Gąsior, Jakub S.; Łaba, Jonasz; Korzeniewski, Kacper; Młyńczak, Marcel (2023) Evaluation of the performance of GPT-3.5 and GPT-4 on the Polish Medical Final Examination. Scientific Reports 13(1), 20512

Sadeek, Soumik Nafis; Hanaoka, Shinya (2023) Assessment of text-generated supply chain risks considering news and social media during disruptive events. Social Network Analysis and Mining 13(1), 96

Shahbaz, Muhammad Saeed; RM Rasi, Raja Zuraidah; Bin Ahmad, MD Fauzi (2019) A novel classification of supply chain risks: Scale development and validation. Journal of Industrial Management and Engineering 12(1), 201–218

Shahsavari, Maryam; Hussain, Omar Khadeer; Saberi, Morteza; Sharma, Pankaj (2024a) Empowering Supply chains Resilience: LLMs-Powered BN for Proactive Supply Chain Risk Identification. Third International Workshop on Linked Data-driven Resilience Research

Shahsavari, Maryam; Hussain, Omar Khadeer; Saberi, Morteza; Sharma, Pankaj (2024b) Event Identification for Supply Chain Risk Management Through News Analysis by Using Large Language Models. The Review of Socionetwork Strategies 18(2), 255–278

Shishehgarkhaneh, Milad Baghalzadeh; Moehler, Robert C.; Fang, Yihai; Hijazi, Amer A.; Aboutorab, Hamed (2024a) Transformer-based Named Entity Recognition in Construction Supply Chain Risk Management in Australia. IEEE Access 12, 41829–41851

Shishehgarkhaneh, Milad Baghalzadeh; Chan, Melissa; Moehler, Robert C.; Fang, Yihai; Hijazi, Amer A.; Aboutorab, Hamed (2024b) Named Entity Recognition in Construction Supply Chain Risk Management using Transformer-Based Models and Genetic Algorithm-Driven Hyperparameter Tuning. International Conference on Electrical, Computer and Energy Technologies (ICECET 2024)

Singh, Sushant; Mahmood, Ausif (2021) The NLP Cookbook: Modern Recipes for Transformer Based Deep Learning Architectures. IEEE Access 9, 68675–68702

Sokolova, Marina (2018) Big Text advantages and challenges: classification perspective. International Journal of Data Science and Analytics 5(1), 1–10

Soliman, Ghada; Zaki, Hozaifa; Kilany, Mohamed (2025) A comparative analysis of encoder only and decoder only models for challenging LLM-generated STEM MCQs using a self-evaluation approach. Natural Language Processing Journal 10, 100131

Srivastava, Mohit; Rogers, Helen (2022) Managing global supply chain risks: effects of the industry sector. International Journal of Logistics 25(7), 1091–1114

Strimel, Grant P; Xie, Yi; King, Brian; Radfar, Martin; Rastrow, Ariya; Mouchtaris, Athanasios (2023) Lookahead When It Matters: Adaptive Non-causal Transformers for Streaming Neural Transducers. Proceedings of the 40th International Conference on Machine Learning, 1–23

Sun, Jun; Wen, Xin; Ping, Gang; Zhang, Mingxuan (2024) Application of News Analysis Based on Large Language Models in Supply Chain Risk Prediction. Journal of Computer Technology and Applied Mathematics 1(3), 55–65

Surampudi, Yeswanth (2024) Big Data Meets LLMs: A New Era of Incident Monitoring. Volume 1. Libertatem Media Private Limited, Ahmedabad, India

Tummala, Rao; Schoenherr, Tobias (2011) Assessing and managing risks using the Supply Chain Risk Management Process (SCRMP). Supply Chain Management 16(6), 474–483

Varma, T. Nn; Khan, D. A. (2014) Information Technology in Supply Chain Management. Journal of Supply Chain Management Systems 3(3), 35–46

Vaswani, Ashish; Shazeer, Noam; Parmar, Niki; Uszkoreit, Jakob; Jones, Llion; Gomez, Aidan N.; Łukasz, Kaiser; Polosukhin, Illia (2017) Attention Is All You Need. 31st Conference on Neural Information Processing Systems, 6000–6010

Vilko, Jyri P.P.; Hallikas, Jukka M. (2012) Risk assessment in multimodal supply chains. International Journal of Production Economics 140(2), 586–595

Vlachos, Ilias; Reddy, Pulagam Gautam (2025) Machine learning in supply chain management: systematic literature review and future research agenda. International Journal of Production Research, 1–30

Wang, Yan-Ling (2012) Design and operating for the logistics systems. Advanced Materials Research 433;440, 3101–3105

Wang, Yue; Le, Hung; Gotmare, Akhilesh Deepak; Bui, Nghi D. Q; Li, Junnan; Hoi, Steven C. H (2023) CodeT5+: Open Code Large Language Models for Code Understanding and Generation. Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing, 1069–1088

Wang, Yiming; Zhang, Ziyang; Chen, Hanwei; Shen, Huayi (2024) Reasoning with Large Language Models on Graph Tasks: The Influence of Temperature. 5th International Conference on Computer Engineering and Application, 630–634

Xia, Long; Shen, Wenqi; Fan, Weiguo; Wang, G. Alan (2024) Knowledge-Aware Learning Framework Based on Schema Theory to Complement Large Learning Models. Journal of Management Information Systems 41(2), 453–486

Xia, Peipei; Zhang, Li; Li, Fanzhang (2015) Learning similarity with cosine similarity ensemble. Information Sciences 307, 39–52

Xie, Yue; Zhou, Shenghan; Chang, Wenbing; Zhao, Jun (2016) An improved supplier selection model for equipment R&D project with independent fuzzy cost information. 28th Chinese Control and Decision Conference, 6502–6505

Xu, Xinhan; Chen, Xiangfeng; Jia, Fu; Brown, Steve; Gong, Yu; Xu, Yifan (2018) Supply chain finance: A systematic literature review and bibliometric analysis. International Journal of Production Economics 204, 160–173

Yamaguchi, Atsuki; Ozaki, Hiroaki; Morishita, Terufumi; Morio, Gaku; Sogawa, Yasuhiro (2023) How does the task complexity of masked pretraining objectives affect downstream performance?. Findings of the Association for Computational Linguistics: ACL 2023, 10527–10537

Yan, Zini; Liang, Hongyu; Wang, Jingya; Zhang, Hongbin; da Silva, Alisson Kwiatkowski; Liang, Shiyu; Rao, Ziyuan; Zeng, Xiaoqin (2024) PDGPT: A large language model for acquiring phase diagram information in magnesium alloys. Materials Genome Engineering Advances 2(4)

Yang, Jingfeng; Jin, Hongye; Tang, Ruixiang; Han, Xiaotian; Feng, Qizhang; Jiang, Haoming; Zhong, Shaochen; Yin, Bing; Hu, Xia (2024) Harnessing the Power of LLMs in Practice: A Survey on ChatGPT and Beyond. ACM Transactions on Knowledge Discovery from Data 18(6), 160

Yao, Shunyu; Hu, Jie; Sun, Chuxiong; Gao, Zhiqiao; Liu, Ning (2022) Key Phrase Extraction based on Pre-trained Language Models. Proceedings of the 6th International Conference on Electronic Information Technology and Computer Engineering, 941–945

Yao, Yifan; Duan, Jinhao; Xu, Kaidi; Cai, Yuanfang; Sun, Zhibo; Zhang, Yue (2024) A survey on large language model (LLM) security and privacy: The Good, The Bad, and The Ugly. High-Confidence Computing 4(2), 100211

Zand, Fardad; Solaimani, Sam; van Beers, Cees (2015) A Role-based Typology of Information Technology: Model Development and Assessment. Information Systems Management 32(2), 119–135

Zhang, Justin Zuopeng; Gupta, Shivam; Srivastava, Santosh Kumar; Routray, Susmi; Bag, Suraji (2024) Exploring the Potential of Large Language Models in Supply Chain Management: A Study Using Big Data. Journal of Global Information Management

Zhang, Yunyi; Zhong, Ming; Ouyang, Siru; Jiao, Yizhu; Zhou, Sizhe; Ding, Linyi; Han, Jiawei (2024) Automated Mining of Structured Knowledge from Text in the Era of Large Language Models. Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 6644–6654

Zhao, Lanxin; Gao, Wanrong; Fang, Jianbin (2024) Optimizing Large Language Models on Multi-Core CPUs: A Case Study of the BERT Model. Applied Sciences 14(6), 2364

Zhao, Ming; Hussain, Omar; Zhang, Yu; Saberi, Morteza (2024a) Optimizing Supply Chain Risk Management: An Integrated Framework Leveraging Large Language Models. 2024 IEEE Conference on Artificial Intelligence, 1057–1062

Zhao, Ming; Hussain, Omar; Zhang, Yu; Saberi, Morteza; Leshob, Abderrahmane (2024b) Enhancing Supply Chain Risk Management with Large Language Models: Software Prototyping and Interactive Visualization. 2024 IEEE International Conference on e-Business Engineering, 284–291

Zhong, Chao; Li, Pengjun; Wang, Jinlong; Xiong, Xiaoyun; Lv, Zhihan; Zhou, Xiaochen; Zhao, Qixin (2025) Enterprise violation risk deduction combining generative AI and event evolution graph. Expert Systems 42(1), n/a