Journal of Computer Science

Review Article

Utilizing Large Language Models in Business Process Management: Applications and Challenges

Mahmoud Boghdady Abdel Wahab, Sherif A. Mazen and Iman M. A. Helal

Faculty of Computers and Artificial Intelligence, Cairo University, Giza, Egypt

Article history
Received: 10-03-2025
Revised: 25-06-2025
Accepted: 12-07-2025

Corresponding Author:
Mahmoud Boghdady Abdel Wahab
Faculty of Computers and Artificial Intelligence, Cairo University, Giza, Egypt
Email: mahmoud.boghdady@gmail.com

Abstract: Large Language Models (LLMs) have recently been used in numerous domains, such as Business Process Management (BPM), which has significantly advanced. With LLMs' ability to understand language, reason, and tackle new challenges with minimal guidance, they offer an exciting opportunity to rethink and improve BPM practices. This systematic literature review examines insights from 42 peer-reviewed studies to understand how LLMs influence different stages of the BPM lifecycle. It sheds light on notable advancements and addresses the challenges that need to be overcome to unlock their full potential. Furthermore, we present an interactive Streamlit application that demonstrates the practical application of LLMs across all five stages of the BPM lifecycle using zero-shot learning, showcasing their potential to automate and enhance BPM tasks. We aim to deepen our understanding of the impact of LLMs on the evolution of BPM practices through a thorough review of current applications and future possibilities. The selected research papers cover LLM representation in various domains: process modeling (14%), process analysis and optimization (14%), process execution and monitoring (11%), process mining (19%), and generic capabilities and challenges (42%). Our findings underscore the growing importance of LLMs in addressing complex BPM scenarios while raising critical questions about scalability, interpretability, and fairness. Finally, this paper presents the technical, ethical, and practical challenges of integrating LLMs into BPM environments.

Keywords: Large Language Models, Business Process Management, Process Mining, Process Modeling, Process Optimization

Introduction

The progress made in artificial intelligence in the last several years with the creation of Large Language Models (LLMs) has been enormous, with applications ranging from healthcare to finance, and beyond. LLMs have demonstrated disruptive potential in the field of Business Process Management (BPM) by automating operations that were previously performed manually or via semi-automated methodologies. Models, such as OpenAI's GPT series, stand out for their advanced ability to read and generate text that closely matches human speech. These qualities have allowed LLMs to improve numerous BPM operations, including process modeling and real-time monitoring, resulting in increased efficiency, accessibility, and correctness.

To gain a full understanding of LLMs' incorporation of LLMs into the BPM, this survey employs a rigorous approach that includes an intensive evaluation of over 60 publications published between 2020 and 2024. These materials were discovered by searching academic databases such as Google Scholar and Connected Papers for keywords such as "large language models in business process management," "LLM," "BPM," "process mining," and "GPT."

From the original pool, 42 publications were chosen for in-depth research based on their relevance to BPM activities across the process lifecycle phases: process modeling and discovery, analysis, optimization, execution, and monitoring. The findings show that LLMs provide considerable benefits throughout several BPM stages. LLMs simplify the BPM for non-expert users by automating manual activities, such as creating process models from text descriptions. They also minimize resource use (Teubner *et al.*, 2023; Grohs *et al.*, 2024).

To the best of our knowledge, this is the first study to provide a comprehensive review of LLMs' applications across the entire BPM lifecycle, coupled with an

DOI: 10.3844/jcssp.2025.1921.1932

interactive tool that allows users to experience these applications firsthand.

LLMs enable a deeper process analysis, helping businesses identify inefficiencies, bottlenecks, and opportunities for change. They provide simple natural language interfaces for monitoring and execution, enabling users to control and analyze processes in real-time, thus boosting their decision-making abilities. However, there are some challenges to integrating LLMs into BPM. Such as data privacy and security, in the case of the sensitive nature of business data that these models may manage. Additionally, LLMs are susceptible to "hallucination," which causes them to provide erroneous outputs that, if ignored, may influence process choices. Reliability and ethical issues are considered major challenges as the use of AI in decision-making processes. Furthermore, integrating LLMs with the current BPM infrastructure may require considerable architectural modification. These results imply that although LLMs have the potential to revolutionize BPM, their use must be carefully controlled to protect data security and process integrity.

To solve these issues, it is advised to customize a specific LLM to BPM and improve the dependence of models. This could be done by integrating knowledge graphs or improved prompts and developing privacy-preserving methods. The integration of LLMs into BPM has greatly improved companies' performance and optimized their processes. One of the benefits is automating repetitive tasks and enhancing process visibility. LLMs enable more efficient and intuitive BPM architecture. However, when combined, there are problems with model reliability, data security, and ethical considerations. Moving forward, research should focus on creating domain-specific LLMs tailored for BPM, enhancing model interpretability to ensure trustworthy AI-driven processes, and establishing secure data handling practices.

As BPM technology evolves, collaboration between researchers and industry practitioners will be essential for refining and responsibly applying these AI-driven models, supporting more adaptive and robust BPM systems.

Large Language Models (LLMs)

Significant progress has been made in Artificial Intelligence (AI) and Natural Language Processing (NLP) owing to LLMs, which have been greatly enhanced by advancements in deep learning, particularly the transformer model introduced by Vaswani *et al.* (2017). LLMs are very important as they can understand and handle complex natural language, generate human-like responses, and handle complex tasks such as text generation, translation, automatic summarization, and answering questions. They can be used in many different fields, including healthcare, education, finance, and technology.

According to Kojima *et al.* (2022), LLMs can be quite effective as zero-shot problem solvers. They can address complex issues and find solutions without requiring specific instructions. They can understand and generate human-like text, which is useful for tasks such as documenting processes and creating natural language interfaces in BPM (Vidgof *et al.*, 2023). They can also handle complex reasoning, draw logical conclusions, and connect different pieces of information (Nori *et al.*, 2023).

However, LLMs have certain limitations and challenges. They may struggle with certain tasks and may not always provide accurate results. In addition, they require considerable computing power and data, which can be logistical hurdles. It is important to be aware of these limitations to address the upcoming challenges and get the most out of them. For example, LLMs can make errors or "hallucinate" information, especially on topics on which they were not specifically trained. They might also reflect or even amplify the biases present in their training data, which can lead to unfair outcomes if not handled carefully.

Data security and the risk of exploitation, particularly with sensitive business information, present other serious concerns (Teubner *et al.*, 2023). Despite their impressive capabilities, LLMs do not truly understand how humans think, and their responses are based on patterns in their training data. It is crucial to understand their strengths and weaknesses to use LLMs effectively in areas such as BPM. Accuracy, reliability, and ethics are the key considerations when integrating LLMs into these domains. Knowing what these models can or cannot do will help ensure they meet the standards for accuracy, reliability, and ethical use.

Business Process Management (BPM)

Business Process Management (BPM) concepts ensure that an organization works more efficiently and provides flexibility. They involve using different methods and tools to design, analyze, carry out, monitor, and improve the processes. The goal is to streamline operations, spot areas that need to be fixed, and boost overall performance. Essentially, BPM aims to help organizations run smoothly while adapting to changes in their business environment (Dumas *et al.*, 2018). The BPM lifecycle consists of five main steps: design, modeling, execution, monitoring, and optimization. The first step is the design phase, in which teams determine what the processes are, set performance goals, and create rules. In the modeling phase, the team visually draws out these processes and checks their relations. During execution, the team implements the processes and monitors how things work. Finally, in the optimization phase, the team used monitoring data to improve the process. Therefore, organizations continually refine and adapt their workflow. In the past, adopting BPM concepts meant considerable manual work. Teams must spend a lot of time and effort on their tasks. This adoption requires significant professional expertise and could lead to slow, costly, and error-prone results. Therefore, the traditional BPM methods can be rigid and inefficient.

Today, there is a growing interest in using automation and advanced analytics to improve BPM. This shift aims to reduce manual tasks and boost both efficiency and accuracy (Beerepoot *et al.*, 2023). Leveraging innovative technology can overcome the limitations of older BPM methods and make the processes more flexible and effective. In today's rapidly changing business environment, BPM faces several challenges.

Complex processes are difficult to model and analyze, and companies need to stay adaptable to keep up with market changes, customer needs, and regulations. The increasing amount and variety of data add another layer of complexity to the traditional methods (Loyola-González, 2023). Integrating BPM with other systems and new technologies remains challenging. Moreover, employees can face resistance if the required changes are too disruptive or complicated for their current work routines (Dumas *et al.*, 2018; Grisold *et al.*, 2022).

Materials and Methods

Integrating LLMs into the BPM can be a game changer to address some of the old challenges in the BPM. LLMs have powerful language-processing and reasoning abilities that can automate and improve many parts of the BPM process. This can change how organizations manage their workflows, leading to better productivity and new ideas (Vidgof *et al.*, 2023). LLMs are effective at handling large amounts of text related to business processes.

They can help companies to better understand data and make more informed decisions. For example, LLMs can analyze customer feedback to find patterns, extract important information from contracts, and even automate the creation of process documents. As LLM technology continues to advance, it is likely to have a significant impact on how business processes are managed and improved. Non-experts can benefit from using LLMs to create process models by automating generation from natural language (Grohs *et al.*, 2024; Kourani *et al.*, 2025; 2024a; Berti *et al.*, 2024a). In addition, LLMs facilitate the analysis of process models and execution data, offering insights and suggestions for optimization (Berti *et al.*, 2024b).

In addition, they provide natural language interfaces that make process monitoring, control, and real-time anomaly detection more intuitive (Vidgof *et al.*, 2023). They also enhance event-log analysis, process discovery, and compliance checking in process mining (Berti *et al.*, 2024a). However, introducing Language Learning Models (LLMs) into Business Process Management (BPM) is challenging. Teubner *et al.* (2023) points out that we need to carefully think about possible issues like data privacy, how reliable the models are, and any ethical concerns about using LLMs effectively. When we examine how LLMs can fit into BPM, it is important to consider both the benefits and potential hurdles. By addressing these challenges, organizations can create the most LLMs while managing risks and ethical issues. It is all about finding the right balance to effectively and responsibly utilize LLMs in BPM.

Figure 1 illustrates how LLMs are now being utilized across various business domains, where they can complement BPM to enhance overall process effectiveness. In general, a large language model is the central part of the BPM. Every element plays an important role in incorporating LLMs into BPM Systems to improve efficiency, help automation, and secure decision-making.

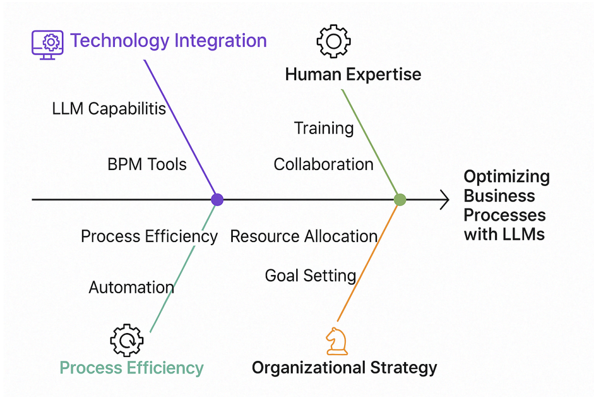

Fig. 1: Relationship of LLMs in different business domains

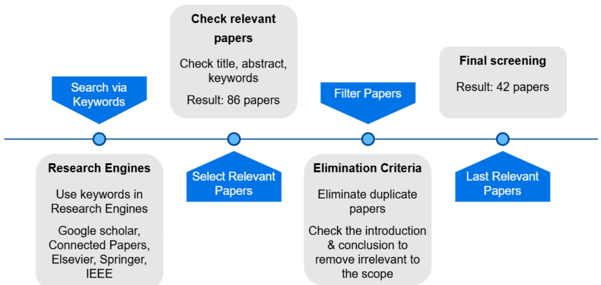

Fig. 2: Selection methodology for relevant research papers

Survey Details and Settings

Our methodology examined how LLMs are being applied in the BPM. We also sought to identify the challenges that businesses face in process management and how LLM technology can address these challenges. This methodology, illustrated in Figure 2, allows us to connect technological innovations with real-world problems in BPM. We used a set of carefully selected keywords. Terms like "Large Language Models in Business Process Management," "LLM," "BPM," "Process Mining," "Process Challenges," "Technological Barriers in Process Management" and other related keywords (Figure 3) were combined with Boolean operators to create a comprehensive search strategy. This helped us find studies that not only discussed LLM applications but also highlighted the challenges businesses face in managing their processes.

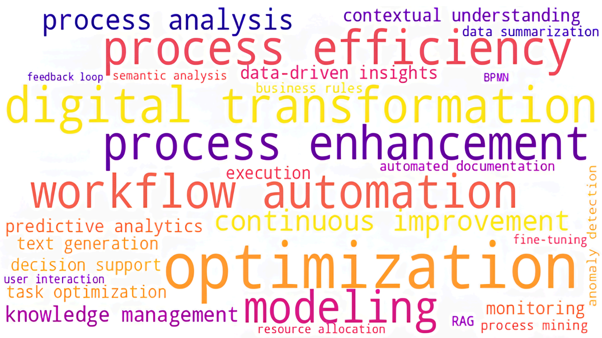

Fig. 3: Research keywords for LLMs in BPM lifecycle

For example, we utilized Google Scholar search terms such as "LLM-based process automation in the BPM lifecycle" and "GPT for business process optimization" to obtain pertinent literature.

At each phase of the review, we clearly stated the research questions and corresponding methods. Each phase explained how LLMs are applied to address specific BPM issues. To ensure a proper survey, a systematic selection of LLM applications in the BPM was conducted. Several selection criteria were used to identify the relevant studies. First, only studies explicitly highlighting the relationship between BPM and LLM applications were included. This ensured that the papers selected were closely related to the survey topic on BPM innovations using LLM technologies. To ensure high-quality studies, we prioritize studies published in peer-reviewed journals, conferences, and workshops. This rule guarantees the scientific quality of literature. Given the rapid advancements in LLM technologies, this study relied on recent papers dating to the last five years to examine recent developments and emerging trends. The survey aimed to include studies highlighting many BPM applications, such as automation, optimization, and decision support. The chosen studies used several methods, such as experiments, case studies, and analyses, which helped us gain a broader understanding of LLM.

Our study revealed that research papers presenting in-depth methodologies, technical models, or specific case deployments applicable to the domain were especially useful as they increased the transparency, applicability, and reproducibility of the results. Following these criteria, the selected studies provide a complete and current overview of the applications of LLMs in BPM.

The methodology used was found on sample questions such as "LLM for BPM" and "LLM in process modeling". We ensured a strict selection process. We included only peer-reviewed articles that were relevant to LLMs in the BPM, with a solid theoretical or practical contribution, and were published between 2020 and 2025. We excluded papers that were not in English, were duplicates, lacked a strong methodology, or had an ambiguous connection to LLM technology. The screening process consisted of three stages. First, we looked at the titles and abstracts to identify papers that met our criteria. We then reviewed the full text of each paper to evaluate its methodology, contributions, and potential impacts. In the final stage, we prioritized studies that offered fresh insights into BPM challenges and explored how LLMs might offer solutions. Out of an initial set of 86 papers, we selected 42 for in-depth analysis. Table 1 shows their distribution: 15 papers from Google Scholar, one paper from Elsevier, 13 papers from Springer, 4 papers from IEEE, and 9 papers from Connected Papers. This mix of sources helped us balance foundational research with applied studies, giving us a well-rounded understanding of both the BPM challenges and the potential role of LLMs.

Table 1: Relevant research papers 2020-2025.

Sources	Number of Papers	Number of Selected Papers
Google Scholar	36	15
Elsevier	1	1
Springer	13	13
IEEE	4	4
Connected Papers	32	9
Total	86	42

LLMs in Process Modeling and Enhancement

Automated Process Model Generation LLMs are making a big impact on applying BPM by automating the creation of process models from text. This advancement is making process modeling more accessible and efficient for many people within organizations, which can lead to increased innovation and adaptability. Grohs *et al.* (2024) discovered that LLMs like GPT-4 can turn natural language descriptions into executable process models, such as BPMN and Petri nets. Their findings highlight how LLMs can manage various BPM tasks effectively, providing valuable insights and suggesting new possibilities for the future.

Kourani *et al.* (2024b) introduced ProMoAI, a tool that uses LLMs to automatically generate process models from text. ProMoAI includes features like advanced prompt engineering and the use of the Partially Ordered Workflow Language (POWL). It also offers a secure environment for running Python code, handles errors with LLM self-refinement, and provides interactive feedback. Additionally, it can convert models into standard formats like BPMN and Petri nets. This tool shows how LLMs can make process modeling simpler and more user-friendly.

In another research, Neuberger *et al.* (2025) proposed a universal strategy for extracting process model information from text. These developments illustrate

DOI: 10.3844/jcssp.2025.1921.1932

how LLMs are advancing and streamlining the process modeling phase in BPM.

Process Information Extraction LLMs are not just about creating detailed process models, they are also skilled at extracting useful information from unstructured data. This is especially helpful when processes are not well-documented or are a bit unclear.

Rizk *et al.* (2024) explored how LLMs can manage various types of data in process mining, such as graphs, sequences, metadata, and unstructured content. They developed models designed specifically for business process data and suggested using a Mixture-of-Modality-Experts (MoME) Transformer architecture to effectively handle different data types. This highlights how LLMs can significantly enhance process analysis and optimization.

De Michele *et al.* (2025) developed a framework that leverages Large Language Models (LLMs) for the automatic extraction of process models from unstructured text descriptions. The study demonstrates the ability of LLMs to effectively identify process elements, activities, and relationships from natural language and consequently produce structured BPMN models automatically, independent of human involvement.

Enhancement of Existing Process Models LLMs can bring a lot of value to process models by adding important context and details that make them more understandable and complete. This documentation improvement can lead to better operational efficiency.

Zhu *et al.* (2023) showed how LLMs could transform simple process diagrams into detailed textual descriptions. They used LLMs to transform Conditional Process Trees (CPTs) into Business Process Text Sketches (BPTSs), which show how these models can enhance the clarity and detail of process documentation. The authors discuss the power of LLMs in addressing common challenges associated with this task. On the other hand, the work demonstrates a unique use of LLM with divide-and-conquer methodologies to break down and resolve complicated hierarchical PTs. This highlights how LLMs may effectively help with process modeling and transformation. Demonstrating how LLMs may improve data production and precision in BPM. Overcome Dataset Scarcity: The approach also tackles a major constraint in process modeling insufficient datasets by giving a method for creating large-scale datasets, which can considerably improve process model extraction (PME).

Vu *et al.* (2026) examine the idea of Agentic Business Process Management (ABPM), focusing on the way autonomous software agents meet business process management. The authors interviewed 22 BPM practitioners to discover their expectations, challenges, and potential regarding the application of autonomous agents in business processes. The paper presents a technology-agnostic definition of ABPM addressing both "agents for BPM" (utilizing agents to improve business processes) and "BPM for agents" (managing autonomous agents in organizational processes). Their findings present practitioners' anticipated benefits (efficiency improvement, predictive insights) and worries (bias, lack of transparency, job displacement).

LLMs in Process Analysis and Optimization

Automated Process Analysis LLMs are making a significant impact on process analysis and optimization. They can analyze process and execution data, helping to pinpoint inefficiencies and offer suggestions for improvement. For instance, Berti *et al.* (2024b) demonstrated how LLMs can generate insightful hypotheses and queries. They introduced methods for simplifying event logs and process models into clearer text. By integrating these techniques into the pm4py process mining library. The authors also achieve optimization by developing and refining effective prompting strategies. By customizing prompts to align with the abstracted process mining information, they enhance LLMs' ability to provide accurate and relevant responses to natural language queries. This tailored approach improves the synergy between process mining and the capabilities of LLMs.

Intelligent Decision Support LLMs have the potential to greatly enhance decision-making in process management. GPT-4, for instance, has demonstrated strong abilities in complex reasoning and has performed well in various benchmarks, including medical examinations. Although much of this research is focused on the medical field, there are promising implications for BPM. GPT-4's capabilities in analyzing complex scenarios, identifying improvement opportunities, explaining anomalies, and evaluating alternative process designs could be highly beneficial for process management and optimization.

Studies by Kampik *et al.* (2024) and Kourani *et al.* (2024) show that LLMs can make process models more understandable, which can lead to more effective analysis and optimization. Their research aims to enhance the interpretability of complex models. Additionally, Fahland *et al.* (2025) investigated how LLMs can clarify business processes, providing valuable insights for process analysis and interpretation. These studies highlight the growing role of LLMs in enhancing decision support and analytical capabilities within BPM.

Feli *et al.* (2025) introduces an LLM-based agent for analyzing physiological data, and more particularly, Heart Rate (HR) estimation using PPG (PhotoPlethysmoGram) signals. Feng *et al.* (2025) created ANESBENCH, a cross-lingual benchmark for assessing LLM reasoning in anesthesiology at three cognitive levels: System 1 (fact retrieval), System 1.x (hybrid reasoning), and System 2 (decision-making with complexity). By thoroughly assessing more than 30

LLMs, they determined the major factors influencing performance, which include model size, Chain of Thought length, and language transferability.

LLMs in Process Execution and Monitoring

Natural Language Interfaces for Process Control Integrating LLMs into BPM's execution and monitoring phases offers exciting possibilities for enhancing real-time decision-making and making BPM systems more accessible. Vidgof *et al.* (2023) highlight how LLMs can transform process management by interpreting natural language queries, providing clear explanations of process behaviors and anomalies, and offering conversational interfaces for process control. This could make BPM systems more intuitive for non-experts' improvement.

Real-Time Process Monitoring and Anomaly Detection LLMs have the potential to revolutionize the way we handle and analyze big datasets in real time. It can help a system administrator to look out for issues in the processes, alerting business owners right when something goes wrong and breaking it down in simple terms. Dumas *et al.* (2018) use AI-enhanced Business Process Management Systems (ABPMSs) to keep an eye on things and adjust on the fly. Meanwhile, Beheshti *et al.* (2023) have come up with ProcessGPT, a framework that uses generative AI to improve decision-making throughout the BPM lifecycle. It provides timely insights and automates tasks, making process execution and monitoring much smoother and more effective.

Derakhshan *et al.* (2025) examine how GPT-4o can detect rework anomalies in business processes. The researchers developed a tool that translates event logs into structured representations and identifies unnecessarily repeated processes across several anomaly distributions.

Supply Chain Risk Management LLMs can make a big difference in how we handle and monitor processes, especially in managing supply chain risks. Zhao *et al.* (2024) introduce an innovative framework that leverages LLMs to improve how we manage these risks. Their system pulls data from news sources and supplier databases to spot potential risk events. They have also created a system for labeling these risks using established categories, and they use semantic analysis to match the risks with the right labels. This approach highlights how LLMs can enhance decision-making by providing clearer, more timely insights into risks, even in the fast-paced world of supply chains.

LLMs in Process Mining

Enhanced Event Log Analysis LLMs are changing the game when it comes to analyzing event logs. They can turn complex log data into easy-to-understand explanations, making process mining more accessible for non-experts and giving deeper insight into how processes work.

Few Research papers by Berti *et al.* (2024a) and Jessen *et al.* (2023) show that LLMs can handle a lot of heavy lifting in process mining. They are great at breaking down event logs, spotting patterns and anomalies, and answering questions about process performance. Kermani, Seddighi, and Maghsoudi (2024) took this a step further by integrating ChatGPT with process mining, which improved user experience through better prompt engineering. This integration makes process mining not just more accessible, but also more user-friendly. Additionally, Berti *et al.* (2025), Berti and Qafari (2023) have provided evaluations of LLMs in this field.

Intelligent Query Generation LLMs can significantly enhance process mining by making it easier to formulate and understand complex queries about process behaviors. This improvement lowers the barrier to using process mining tools and allows for a more thorough analysis of process data.

Busch *et al.* (2023) investigated how LLMs can be applied to process mining. Their research highlighted several advantages: LLMs can manage limited data effectively, offer natural language interactions, and optimize inputs through tailored prompts. They found that well-designed prompts could improve how users interact with process mining tools, making it easier to ask detailed questions and gain deeper insights.

Process Mining Datasets and Benchmarks Researchers are making exciting strides in process mining with the help of LLMs, and they are doing so by developing new datasets and benchmarks. For example, Sola *et al.* (2023) have worked with the SAP Signavio Academic Models, a massive collection of over 1 million business models in BPMN notation. This resource is crucial for training and evaluating LLMs on process-related tasks. Similarly, Li *et al.* (2023) have created the MaD dataset, which focuses on extracting process information from natural language descriptions in interviews. These resources are important for advancing LLMs in process mining, giving us standardized data to test and refine new approaches.

LLMs Case Studies

In our study, we demonstrated that it is possible to utilize large language models to effectively simplify business process management across all phases, from modeling and analysis to execution, mining, and risk management. With advanced prompting techniques and workflow-specific languages, authors showed how process diagrams can be automatically generated and optimized by such models with superior results compared to traditional tools without needing heavy tuning. Integrations with popular libraries enabled automated hypothesis generation, event-log simplification, and anomaly detection, with significant reductions to root-cause analysis and resolution times.

In real-world pilots, fine-tuned generative transformers e.g., ProcessGPT produced executable flows and assisted real-time monitoring, reducing mean time to resolution significantly. Also, GPT-4o-based tools achieved near-perfect accuracy in one-shot and few-shot detection of rework anomalies, while conversational prompt strategies dramatically enhanced usability in enterprise settings. In process mining, "Chit-Chat or Deep Talk" prompt engineering improved conversational interfaces' usability by 72 % in a 17-company pilot, while the PM-LLM-Benchmark offers the first domain-specific evaluation of LLMs on mining tasks, showing commercial models' strength but also open-source limitations and evaluation biases. Collectively, our findings point to the transformative potential of LLMs in BPM and the necessity of custom prompts, domain-specific fine-tuning, thorough evaluation, and strategies to ensure data privacy, prevent hallucinations, and enable explainability for safe, enterprise-grade deployment.

In addition to the case studies discussed, we have developed an interactive Streamlit application that leverages LLMs to automate and enhance the five key stages of BPM: Identification, Discovery, Analysis, Redesign, and Monitoring. This tool enables users to input business process descriptions and receive automated analyses, visualizations, and BPMN models, demonstrating the practical utility of LLMs in BPM without the need for custom training. The application serves as proof of concept, illustrating how LLMs can make BPM more accessible and efficient for both experts and non-experts alike. For a live demonstration, the application can be accessed at https://llm-powered-bpm-lifecycle.streamlit.app/ .

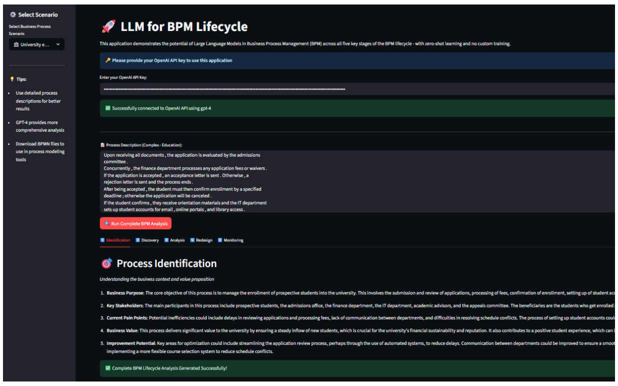

Fig. 4: LLM powered BPM lifecycle Streamlit Application

Results and Discussion

Current Research Utilizing LLMs in BPM Lifecycle

We examined insights from 42 peer-reviewed studies to understand how LLMs are influencing different stages of the BPM lifecycle. It sheds light on notable advancements while also addressing the challenges that need to be overcome to unlock their full potential. By presenting a thorough review of current applications and future possibilities, we aim to deepen the understanding of LLMs' impact on the evolution of BPM practices. The research covers process modeling with 6 papers (14%), process analysis and optimization with 6 papers (14%), process execution and monitoring with 5 papers (11%), process mining and datasets with 8 papers (19%), and generic capabilities and challenges with 17 papers (42%). These findings highlight the breadth of exploration across various BPM aspects and emphasize the growing significance of LLM applications in enhancing BPM outcomes.

With LLMs, organizations can make these processes faster and more efficient, saving both time and resources. This cutting-edge technology opens new possibilities for businesses to optimize their operations and keep up with the fast-paced changes in today's market. Capturing inefficiencies, bottlenecks, and areas for improvement in business processes is a key part of managing and optimizing them. By incorporating LLMs into this, we can significantly boost the depth, speed, and accessibility of process analysis and optimization. This means quicker insights and more effective improvements, making the entire process management more efficient.

Integrating LLMs into the execution and monitoring phases of BPM could significantly improve real-time decision support and process control. These models have the potential to provide more intuitive interfaces for process stakeholders, enhancing the overall user experience. While research in this area is still developing, the potential benefits suggest a promising future for making process management systems more efficient and accessible.

Process mining is like a bridge between data mining and business process management. It has a lot of potential for making sense of event logs, improving how we uncover process details, and making process mining tools easier to use. Using LLMs can help enhance this process. By analyzing event logs, LLMs can help us get deeper insights and make the whole process of understanding and refining real-world processes much more efficient and user-friendly.

Challenges and Limitations

Bringing LLMs into BPM is a promising move, offering a range of potential benefits. But it also comes with its own set of challenges and limitations that need to be carefully considered. The literature highlights several key issues, from technical hurdles to ethical concerns.

Data Privacy and Security: Integrating LLMs into BPM holds great promise, but it also raises some critical concerns, especially around data security. Teubner *et al.* (2023) emphasize that using LLMs with sensitive or proprietary business data, especially in cloud-based environments, comes with risks. These include potential exposure of confidential information, compliance with

DOI: 10.3844/jcssp.2025.1921.1932

data protection laws, and ensuring data security throughout its use.

Addressing these concerns requires developing robust privacy and security measures tailored specifically for LLMs in BPM contexts. It is essential to guard sensitive data and address potential vulnerabilities.

Model Reliability and Accuracy: Leveraging LLMs in BPM holds great promise but also comes with its share of challenges. One of the key issues is the tendency of LLMs to generate outputs that might sound correct but are wrong; this is known as "hallucination." In BPM, where decisions based on LLM-generated insights can have significant consequences, this is particularly concerning. Nori *et al.* (2023) have highlighted that even advanced models like GPT-4 can sometimes produce information that appears credible but is not accurate.

In the field of software engineering, Hou *et al.* (2023) have reviewed the challenges and opportunities of using LLMs in complex technical areas. Their findings offer valuable insights into how similar problems might affect BPM. Yang *et al.* (2024) proposed using knowledge graphs to improve the factual accuracy of LLMs, which could help address some of these issues.

To effectively integrate LLMs into BPM, it is crucial to tackle these challenges. This involves ensuring that the outputs from LLMs are accurate, addressing any biases present in the training data, and thoroughly validating the results. By focusing on these areas, we can better leverage the benefits of LLMs while minimizing the risk of making erroneous business decisions.

Integration with Existing BPM Systems: Integrating LLMs into current BPM tools and workflows presents a variety of technical challenges. Dumas *et al.* (2018) highlight that to make this integration successful; we need to design new architectures that can smoothly incorporate Artificial Intelligence into existing BPM systems.

To ensure that LLMs are compatible with established process modeling notations and standards. It is also crucial to find the right balance between the adaptability of LLMs and the structured requirements of formal BPM systems. Additionally, creating intuitive interfaces that enable effective collaboration between humans and AI in process management tasks is vital.

Security and Ethical Considerations When we start using AI in decision-making, it is important to think about the ethical implications, including transparency, fairness, and accountability. Teubner *et al.* (2023) emphasize the need for frameworks that ensure AI is used responsibly in BPM systems. One major concern is how to ensure we understand how AI models make their decisions. If we cannot see how a model gives its output, it is hard to trust it. Biases in these models also need to be addressed because they can lead to unfair outcomes or flawed process designs.

This study identifies the common ethical issues and security risks, including the occurrences of LLM hallucinations and reinforcement of biases. In literature, these are presented gradually; certain studies reference them explicitly, while others do not. Nevertheless, through systematically addressing these challenges by creating tailored BPM-specific LLMs or composite models, there is ample scope for raising the level of trust and acceptability.

Insights

This systematic review of the literature consolidates a broad range of current research to produce a set of novel findings that contribute depth to our understanding of LLM incorporation in BPM:

Cross-Stage Integration Though most of the current literature is focused on individual BPM stages, this integration promises significant opportunities for integrating LLMs throughout the end-to-end BPM life cycle, spanning from modeling, analysis, and execution to monitoring and mining. For instance, knowledge from LLM-based process modeling can be utilized in monitoring processes to enable continuous improvement and enhance more dynamic BPM environments.

Enhanced Accessibility and Democratization: LLMs significantly reduced technical barriers to handling processes' work. This review is centered on how natural language-based automated tools allow non-technical stakeholders to actively participate in BPM activities, elevating organizational participation and enhancing the overall performance (e.g., ProMoAI).

By synthesizing these findings, this paper clarifies some opportunities and challenges of LLM-based BPM and provides actionable recommendations to researchers and practitioners for advancing the field.

Open Future Directions

As we look to the future, there are several key areas where research and development could make a significant impact.

Development of BPM-Specific LLMs: Rizk *et al.* (2024) points out the need for LLMs that can effectively handle various types of business process data. They suggest that future research should focus on refining these models with specific knowledge related to process management and exploring multi-modal approaches. Wornow *et al.* (2024) introduced WONDERBREAD, a benchmark that evaluates how well multi-modal models perform in BPM tasks. This is a valuable tool for guiding future research. Similarly, Majumder *et al.* (2024) created DiscoveryBench to assess LLMs on data-driven discovery tasks, which could be very useful in BPM contexts.

Dai *et al.* (2023) showed how ChatGPT can be used to enhance BPM datasets by generating more relevant text data. Lin (2024) worked on improving methods for recognizing structures in PDFs, which is crucial for processing business documents. Eloundou *et al.* (2024) examined how LLMs could impact the job market, providing insights into how BPM roles and skills might evolve. Küsters and van der Aalst (2024) developed a high-performance process mining library with Java and Python bindings in Rust, and there is potential to boost its performance further with LLM coding capabilities.

Busch *et al.* (2023) investigated various prompting strategies for BPM tasks. Looking ahead, several promising research directions are available. One is to develop standardized prompts that address common BPM challenges. Another is to explore dynamic prompting techniques that adapt to different contexts and user needs. Additionally, automating prompts creation and optimization specifically for BPM could be highly beneficial. These research areas offer exciting opportunities to advance how LLMs are used in BPM.

Hybrid Approaches Combining LLMs with Traditional BPM Techniques: Integrating traditional BPM methods with LLMs offers great opportunities. For example, using LLMs with process mining gives us better insights into how processes work and helps us ensure that they are running the way they should. AI-based agents can further enhance this integration by automating the analysis of process data, dynamically adapting to changes, and facilitating seamless collaboration between tools and methodologies. Agents can also assist in breaking down complex process mining tasks into manageable components and integrating deterministic tools with LLM-driven insights for a more comprehensive understanding of processes.

We also might see a lot of improvements by combining LLM outcomes with other formal verification techniques to ensure accuracy and reliability. AI-based agents can orchestrate workflows that combine LLM-generated insights with formal verification methods, ensuring robust and validated results. These agents can continuously refine outputs by learning from verification feedback, enhancing both the accuracy and adaptability of BPM systems over time.

Explainable AI in BPM: In BPM, it is very important to develop explainable AI techniques to make LLM output more understandable and know the reason behind it. This is key for building trust and ensuring accountability for results. Future research could focus on a few key areas: First, we need to create methods that provide clear explanations for the process models and recommendations generated by LLMs. It’s also important to develop visualization tools that help users check the logic and reasoning behind these outputs. Finally, aligning LLM explanations with the specific concepts and terminology used in BPM could greatly improve their relevance and clarity.

Privacy-Preserving Techniques for LLM Usage in BPMWhen working with LLMs and sensitive business data, it is very important to focus on ensuring privacy. We are looking for how to anonymize and encrypt data, so it remains useful for analysis without compromising security. Another promising approach is federated learning, which allows LLMs to learn from data that is spread across different sources, without putting any individual or organization’s privacy at risk. Also, secure multi-party computation techniques, such as those used in blockchain technology, could help keep sensitive data safe during analysis. These strategies are key to leveraging the power of LLMs while ensuring that privacy is always a priority.

Conclusion

The integration of Large Language Models (LLMs) into Business Process Management (BPM) is a major step forward in how we work with processes. This survey tried to cover LLMs' impact on BPM, focusing on process modeling with 6 papers, process analysis and optimization with 6 papers, process execution and monitoring with 5 papers, process mining and datasets with 8 papers, and capabilities and challenges with 17 papers. The survey shows how LLMs can boost efficiency, make systems more accessible, and provide better decision support. Yet, many challenges remain to be tackled, such as ensuring data privacy, improving model reliability and integration, and addressing ethical considerations.

Despite these challenges, our work not only synthesizes the current research on LLMs in BPM but also advances the field by providing a practical, interactive tool, accessible at https://llm-powered-bpm-lifecycle.streamlit.app/ , that showcases the capabilities of LLMs across the entire BPM lifecycle. This tool represents a significant step towards democratizing BPM, making advanced process management techniques accessible to a broader audience.

Moving forward, research areas like automated process model generation and natural language interfaces for BPM systems show great promise. As the field improves, we can expect more sophisticated, domain-specific applications that overcome current limitations and open new possibilities. The future of BPM will likely see cooperation between human expertise and AI-driven insights, with LLMs helping to connect natural language understanding with formal process models. Finally, continuous collaboration between researchers and practitioners will be essential to ensure these technologies effectively address real business needs and challenges and keep improving.

Acknowledgment

The authors would like to thank the anonymous reviewers and colleagues for their insightful feedback and support during the preparation of this manuscript.

Funding Information

This research received no external funding.

Author's Contributions

Mahmoud B. M. Abdelwahab: Contributed to developing the main idea and research methodology for this study, performed the analysis, carried out the investigation work, and wrote the first version of the manuscript.

Sherif A. Mazen: Supervised the research work and helped in checking and confirming the obtained results. Author also contributed to revising and improving the manuscript content.

Iman M. A. Helal: Provided supervision for this research and assisted in verifying the study findings, participated in reviewing the manuscript and suggesting modifications to enhance its quality.

All authors reviewed the final manuscript and agreed on its submission.

Ethics

The authors declare that no ethical issues are anticipated from the publication of this work. This study did not involve human participants, animal subjects, or the use of personal, clinical, or sensitive data. No experiments requiring ethical approval were conducted. The authors further confirm that there are no conflicts of interest to disclose.

References

Beerepoot, I., Di , C., Reijers, H. A., Rinderle-Ma, S., Bandara, W., Burattin, A., Calvanese, D., Chen, T., Cohen, I., Depaire, B., Di Federico, G., Dumas, M., van Dun, C., Fehrer, T., A. Fischer, D., Gal, A., Indulska, M., Isahagian, V., Klinkmüller, C., … Zerbato, F. (2023). The biggest business process management problems to solve before we die. *Computers in Industry*, *146*, 103837. https://doi.org/10.1016/j.compind.2022.103837

Beheshti, A., Yang, J., Sheng, Q. Z., Benatallah, B., Casati, F., Dustdar, S., Nezhad, H. R. M., Zhang, X., & Xue, S. (2023). ProcessGPT: Transforming Business Process Management with Generative Artificial Intelligence. *2023 IEEE International Conference on Web Services (ICWS)*, 546–553. https://doi.org/10.1109/icws60048.2023.00099

Berti, A., Kourani, H., Häfke, H., Li, C.-Y., & Schuster, D. (2024a). Evaluating Large Language Models in Process Mining: Capabilities, Benchmarks, and Evaluation Strategies. *International Conference on Business Process Management (BPM 2024), Held in Eindhoven, The Netherlands*, *511*, 13–21. https://doi.org/10.1007/978-3-031-61007-3_2

Berti, A., Kourani, H., & van der Aalst, W. M. P. (2025). Pm-llm-benchmark: evaluating large language models on process mining tasks pm-llm-benchmark: evaluating large language models on process mining tasks. *International Conference on Process Mining Workshops (ICPM 2024), Held in Lyngby, Denmark*, *553*, 610–623. https://doi.org/10.1007/978-3-031-82225-4_45

Berti, A., & Qafari, M. S. (2023). Leveraging large language models (LLMs) for process mining. *ArXiv*. https://doi.org/10.48550/arXiv.2307.12701

Berti, A., Schuster, D., & van der Aalst, W. M. P. (2024b). Abstractions, Scenarios, and Prompt Definitions for Process Mining with LLMs: A Case Study. *International Conference on Process Mining Workshops (ICPM 2023), Rome, Italy*, *475*, 427–439. https://doi.org/10.1007/978-3-031-50974-2_32

Busch, K., Rochlitzer, A., Sola, D., & Leopold, H. (2023). Just Tell Me: Prompt Engineering in Business Process Management. *BPM 2023 Workshops, Held in Utrecht, The Netherlands*, *468*, 3–11. https://doi.org/10.1007/978-3-031-34241-7_1

Dai, H., Li, Q., Huang, J., Yang, L., Ren, X., & Wang, W. Y. (2023). AugGPT: Leveraging ChatGPT for text data augmentation. *ArXiv Preprint, ArXiv:2302.13007*. https://doi.org/10.48550/arXiv.2302.13007

De Michele, W., Armas Cervantes, A., & Frermann, L. (2025). Automated Business Process Analysis: An LLM-Based Approach to Value Assessment. *International Conference on Advanced Information Systems Engineering (CAiSE 2025)*, *557*, 62–69. https://doi.org/10.1007/978-3-031-94590-8_8

Derakhshan, M., Ceravolo, P., & Mohammadi, F. (2025). Leveraging GPT-4o efficiency for detecting rework anomaly in business processes. *ArXiv Preprint, ArXiv:2502.06918*. https://doi.org/10.48550/arXiv.2502.06918

Dumas, M., La Rosa, M., Mendling, J., & Reijers, H. A. (2018). Process-Aware Information Systems. *International Conference on Artificial Intelligence Applications and Innovations (AIAI 2022), Crete, Greece*, *13223*, 341–369. https://doi.org/10.1007/978-3-662-56509-4_9

Eloundou, T., Manning, S., Mishkin, P., & Rock, D. (2024). GPTs are GPTs: Labor market impact potential of LLMs. *Science*, *384*(6702), 1306–1308. https://doi.org/10.1126/science.adj0998

Fahland, D., Fournier, F., Limonad, L., Skarbovsky, I., & Swevels, A. J. E. (2025). How well can a large language model explain business processes as perceived by users? *Data & Knowledge Engineering*, *157*, 102416. https://doi.org/10.1016/j.datak.2025.102416

Feli, M., Azimi, I., Liljeberg, P., & Rahmani, A. M. (2025). An LLM-powered agent for physiological data analysis: A case study on PPG-based heart rate estimation. *ArXiv:2502.12836*. https://doi.org/10.48550/arXiv.2502.12836

Feng, X., Jiang, W., Wang, Z., Luo, Y., Xu, P., Yu, B., Jin, H., Du, B., & Zhang, J. (2025). AnesBench: Multi-dimensional evaluation of LLM reasoning in anesthesiology. *arXiv:2504.02404*. https://doi.org/10.48550/arXiv.2504.02404

Grohs, M., Abb, L., Elsayed, N., & Rehse, J.-R. (2024). Large Language Models Can Accomplish Business Process Management Tasks. *Business Process Management Workshops*, *492*, 453–465. https://doi.org/10.1007/978-3-031-50974-2_34

Hou, X., Zhang, X., Chen, Y., Wang, H., Li, M., & Li, L. (2023). Large language models for software engineering: A systematic literature review. *ArXiv Preprint ArXiv:2308.10620*. https://doi.org/10.48550/arXiv.2308.10620

Jessen, U., Sroka, M., & Fahland, D. (2023). Chit-chat or deep talk: Prompt engineering for process mining. *ArXiv Preprint ArXiv:2307.09909*. https://doi.org/10.48550/arXiv.2307.09909

Kampik, T., Warmuth, C., Rebmann, A., Agam, R., Egger, L. N. P., Gerber, A., Hoffart, J., Kolk, J., Herzig, P., Decker, G., van der Aa, H., Polyvyanyy, A., Rinderle-Ma, S., Weber, I., & Weidlich, M. (2024). Correction: Large Process Models: A Vision for Business Process Management in the Age of Generative AI. *KI - Künstliche Intelligenz*. https://doi.org/10.1007/s13218-024-00884-3

Kermani, M. A. M. A., Seddighi, H. R., & Maghsoudi, M. (2024). Revolutionizing process mining: A novel architecture for ChatGPT integration and enhanced user experience through optimized prompt engineering. *ArXiv:2405.10689*. https://doi.org/10.48550/arXiv.2405.10689

Kojima, T., Gu, S. S., Reid, M., Matsuo, Y., & Iwasawa, Y. (2022). Large language models are zero-shot reasoners. *Advances in Neural Information Processing Systems*, *35*, 22199–22213. https://doi.org/10.48550/arXiv.2205.11916

Kourani, H., Berti, A., Schuster, D., & Aalst, W. M. P. van dar. (2024). Leveraging large language models for enhanced process model comprehension. *ArXiv:2408.08892*. https://doi.org/10.48550/arXiv.2408.08892

Kourani, H., Berti, A., Schuster, D., & van der Aalst, W. M. P. (2024a). Process Modeling with Large Language Models. *Enterprise, Business-Process and Information Systems Modeling – BPMDS 2024, EMMSAD 2024*, *511*, 229–244. https://doi.org/10.1007/978-3-031-61007-3_18

Kourani, H., Berti, A., Schuster, D., & van der Aalst, W. M. P. (2024b). ProMoAI: Process Modeling with Generative AI. *Proceedings of the Thirty-ThirdInternational Joint Conference on Artificial Intelligence*, 8708–8712. https://doi.org/10.24963/ijcai.2024/1014

Kourani, H., Berti, A., Schuster, D., & van der Aalst, W. M. P. (2025). Evaluating large language models on business process modeling: framework, benchmark, and self-improvement analysis. *Software and Systems Modeling*, *25*(5), 1–23. https://doi.org/10.1007/s10270-025-01318-w

Küsters, A., & Aalst, W. M. P. van. (2024). Developing a high-performance process mining library with Java and Python bindings in Rust. *ArXiv:2401.14149*. https://doi.org/10.48550/arXiv.2401.14149

Li, X., Ni, L., Li, R., Liu, J., & Zhang, M. (2023). MaD: A Dataset for Interview-based BPM in Business Process Management. *2023 International Joint Conference on Neural Networks (IJCNN)*, 1–8. https://doi.org/10.1109/ijcnn54540.2023.10191898

Lin, Daqing. (2024). Revolutionizing retrieval-augmented generation with enhanced PDF structure recognition. *ArXiv*. https://doi.org/10.48550/arXiv.2401.12599

Loyola-González, O. (2023). Process mining: software comparison, trends, and challenges. *International Journal of Data Science and Analytics*, *15*(4), 407–420. https://doi.org/10.1007/s41060-022-00379-0

Majumder, B. P., Xu, Z., Chen, W., Sengupta, B., & Su, I. (2024). DiscoveryBench: Towards data-driven discovery with large language models. *ArXiv*. https://doi.org/10.48550/arXiv.2407.01725

Neuberger, J., Ackermann, L., van der Aa, H., & Jablonski, S. (2025). A Universal Prompting Strategy for Extracting Process Model Information from Natural Language Text Using Large Language Models. *Conceptual Modeling*, *15238*, 38–55. https://doi.org/10.1007/978-3-031-75872-0_3

Nori, H., King, N., McKinney, S. . M., Carignan, D., & Horvitz, E. (2023). Capabilities of GPT-4 on medical challenge problems. *ArXiv*. https://doi.org/10.48550/arXiv.2303.13375

Rizk, Y., Venkateswaran, P., Isahagian, V., Narcomey, A., & Muthusamy, V. (2024). A Case for Business Process-Specific Foundation Models. *Ternational Conference on Business Process Management (BPM 2023)*, *492*, 44–56. https://doi.org/10.1007/978-3-031-50974-2_4

Sola, D., Warmuth, C., Schäfer, B., Badakhshan, P., Rehse, J.-R., & Kampik, T. (2023). *SAP Signavio Academic Models: A Large Process Model Dataset*. *14272*, 453–465. https://doi.org/10.1007/978-3-031-27815-0_33

Teubner, T., Flath, C. M., Weinhardt, C., van der Aalst, W., & Hinz, O. (2023). Welcome to the Era of ChatGPT. *Business & Information Systems Engineering*, *65*(2), 95–101. https://doi.org/10.1007/s12599-023-00795-x

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, Illia. (2017). Attention is all you need. *ArXiv*. https://doi.org/10.48550/arXiv.1706.03762

Vidgof, M., Bachhofner, S., & Mendling, J. (2023). Large Language Models for Business Process Management: Opportunities and Challenges. *Business Process Management Forum*, *490*, 107–123. https://doi.org/10.1007/978-3-031-41623-1_7

Vu, H., Klievtsova, N., Leopold, H., Rinderle-Ma, S., & Kampik, T. (2026). Agentic Business Process Management: Practitioner Perspectives on Agent Governance in Business Processes. *Business Process Management: Responsible BPM Forum, Process Technology Forum, Educators Forum*, *565*, 39–54. https://doi.org/10.1007/978-3-032-02936-2_3

Wornow, M., Zhao, M., Hussain, O., Zhang, Y., & Saberi, M. (2024). WONDERBREAD: A benchmark for evaluating multimodal. *ArXiv*. https://doi.org/10.48550/arXiv.2406.13264

Yang, L., Chen, H., Li, Z., Ding, X., & Wu, X. (2024). Give us the Facts: Enhancing Large Language Models With Knowledge Graphs for Fact-Aware Language Modeling. *IEEE Transactions on Knowledge and Data Engineering*, *36*(7), 3091–3110. https://doi.org/10.1109/tkde.2024.3360454

Zhao, M., Hussain, O., Zhang, Y., & Saberi, M. (2024). Optimizing Supply Chain Risk Management: An Integrated Framework Leveraging Large Language Models. *2024 IEEE Conference on Artificial Intelligence (CAI)*, 1057–1062. https://doi.org/10.1109/cai59869.2024.00192

Zhu, R., Hu, Q., Li, W., Xiao, H., Wang, C., & Zhou, Z. (2023). Business process text sketch automation generation using large language. *ArXiv*. https://doi.org/10.48550/arXiv.2309.01071