Universiteit Leiden
The Netherlands

ICT in Business & the Public Sector

Providing domain knowledge for process mining with ReWOO-based agents

Max W. Vogt
s2098660

Supervisors:
First supervisor: Dr. P.W.H. van der Putten
Second supervisor: Prof. dr. ir. H.A. Reijers (Utrecht University)

MASTER THESIS

Leiden Institute of Advanced Computer Science (LIACS)
www.liacs.leidenuniv.nl

July 9, 2024

Abstract

Process mining practitioners often face the challenge of interpreting complex process data and driving process improvements with limited expertise in process optimization, tools, and the operational context of the organization. This study explores the integration of Large Language Models (LLMs) and LLM-based agentic frameworks into process mining tools to bridge this domain knowledge gap and democratize access to process optimization. By using the knowledge and capabilities of LLMs, we propose a framework that automatically generates and inserts domain knowledge, traditionally supplied by a human domain expert, into process mining tools. Our conceptual vision describes a system capable of executing process mining techniques, such as process discovery, conformance checking, and process enhancement, supported by a conversational interface that guides a user through the process and provides the user with relevant domain knowledge. We developed a Proof-of-Concept (PoC) that leverages a Reasoning WithOut Observation (ReWOO)-based agent to perform process discovery, problem identification, generate ecosystem domain knowledge, and propose potential process improvements. With this approach, we demonstrate the ability of LLM-based systems to produce detailed and domain-relevant insights. This study presents the conceptual vision, the developed proof-of-concept, and the evaluation of the developed system. Our findings suggest that LLM-based systems can deliver meaningful domain knowledge into process mining tools; Challenges such as data availability and system complexity still exist, indicating directions for future work.

Contents

1 Introduction

Process mining offers a series of benefits for organizations, such as providing fact-based insights into processes, alignment with process models, and process improvement [1]. These benefits can, for instance, allow an organization to increase productivity and save costs. However, in the current state of process mining, there are obstacles that organizations often encounter when they try to implement it. Zerbato et al. made an overview of the current process mining practices [2]. One of the issues they found was a lack of domain knowledge, which is often encountered by practitioners of process mining.

Another current issue with process mining and Business Process Management (BPM) is that it can be an inaccessible field for a lot of people. Kampik et al. described it as follows: “Managing business processes is knowledge-intense work, requiring both in-depth expertise with respect to specific tools and skill sets, such as process modeling notations and process data query languages and access to and a good grasp of the knowledge and data that exists about a particular process, typically in a highly complex organizational context. Hence, human BPM experts (individuals or teams) must have a high level of technical and socio-professional skills, as well as substantial experience within a particular organization: the entry bar for successfully running a BPM initiative is high.” [3]. A lack of domain expertise was also identified as one of the challenges within process mining by Martin et al. [4]. Furthermore, Mamudu et al name the presence of subject matter experts and the availability of contextual information as success factors for process mining implementations [5]. This also emphasizes the importance of domain experts and domain expertise for a successful implementation of process mining. Andrews et al. found that process improvements were largely proposed by domain experts and that the process mining tool itself was not able to provide direction for these improvements [6]. They proposed that, in the future, process mining tools should be able to automatically propose improvements without a domain expert.

We propose to explore the potential of using Large Language Models (LLMs) to provide domain knowledge in process mining tools. We conducted expert interviews with a small group of process mining practitioners. These interviews gave an insight into the current state of process mining, the value of domain knowledge within process mining, and potential use cases for LLMs in process mining. We validated the lack of domain knowledge in process mining and the potential of using LLMs to help solve it with these interviews.

Several studies have been done about using LLMs in combination with process mining, such as connecting the open-source process mining tool PM4PY to an LLM [7] [8]. Not only academics are looking into using LLMs in process mining, but we also see interest in this area from commercial companies. For instance, process mining vendors Celonis and Pegasystems are currently developing LLM capabilities in their process mining solutions [9, 10]. However, all of the mentioned efforts are mainly focusing on using LLMs for descriptive tasks in process mining tools, e.g. ‘How long does my process take to run?’ or answering questions regarding process mining expertise, e.g. ’What is conformance checking?’.

This paper proposes a different approach. We want to utilize the capabilities of LLMs to fill the observed gap in domain knowledge in process mining tools. The type of knowledge that would traditionally come from a human domain expert, e.g. knowledge about the sector of the organization. Therefore, our research question is:

‘How can LLM-based agents provide domain knowledge to help users of process mining tools understand and improve processes?’

To study this, we follow the principles of a research by design study as proposed by Peffers et al. [11]. We constructed a conceptual framework that describes a system that can execute process mining techniques and generate domain knowledge. Based on this framework, we developed a Proof-of-Concept (PoC) and evaluated this with a limited set of experiments. The results indicate that LLM-based systems can generate domain knowledge and give directions for future work.

The remainder of this thesis is structured into eight sections. Section 2 discusses the relevant background information for this research, Section 3 presents the related work, and Section 4 describes the methods that were used in this research. In Section 5, we discuss the findings and requirements that we gathered from the interviews. Section 6 presents the design and development of the PoC and Section 7 the results of the evaluation of the PoC. In Section 8 we discuss the limitations of this study and potential directions for future work, and finally in Section 9 we present our conclusions.

2 Background

In this section, we present the theoretical foundation of this study that is necessary to understand the remainder of this study. We discuss three main topics: process mining, LLMs, and technology acceptance. The studies that also dive into using LLMs in process mining (the related work) are not presented in this section but in Section 3.

2.1 Process mining

With this study, we aim to explore the potential added value of using LLMs for supplying domain knowledge to process mining tools. Process mining refers to the technique where data from event logs that are extracted from information systems, are used to discover, monitor, and improve real processes in an organization [12]. Process mining consists of three main techniques: process discovery, conformance checking, and process enhancement. These three techniques fall in the segments of the BPM life-cycle [13].

An event log is a record of sequentially stored events [14], in which each event refers to an activity and relates to a case. An activity is a well-defined part of a process and a case is a process instance. Event logs can also contain additional information, like timestamps for each event, that can also be used in process mining. The three main techniques of process mining (process discovery, conformance checking, and process enhancement) can all make use of an event log, Figure 1 shows the relation between event logs and these three techniques.

We will now briefly explain the three process mining techniques:

1. **Process Discovery:** Process discovery produces a process model based on the data of an event log [15]. These process models can be 'mined' from the event logs using different algorithms, often used algorithms include Alpha miner, Heuristic miner, and Inductive miner [16].

2. **Conformance checking:** Conformance checking involves comparing the process model to reality. Conformance checking does this by comparing actual event data from the process execution to a process model [17]. Using metrics an organization can then generate insights from this comparison, one often used metric is 'fitness'. Fitness states how well the process execution data fits the provided process model.

3. **Process enhancement:** Process enhancement involves improving or extending the existing process model [14]. So process discovery and conformance checking focus on displaying the situation in reality and process enhancement focuses on changing the process to improve it.

2.1.1 Process models

For this research, process models will be used as the visual representation of the processes and as a starting point for process enhancement. Smirnov et al define a business process model as: "key artifacts to represent how work is performed in organizations" [18]. These process models can be used in the organization to for instance support decision making. Process models can be displayed using different modeling standards, often used standards are Petri nets and BPMN [14]. Different standards can be used to display the same business process. Figure 2 and 3 display the same

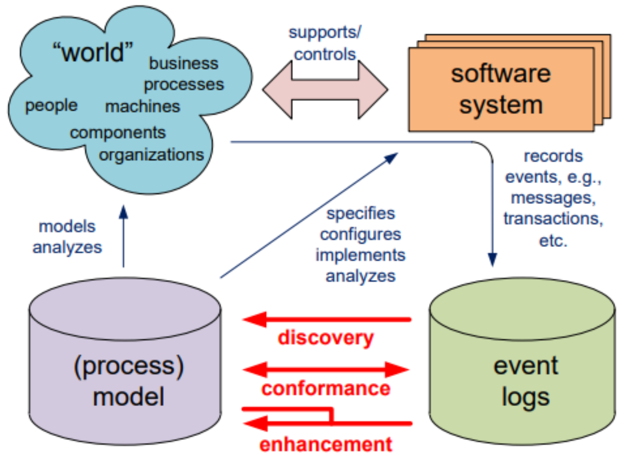


Figure 1: The relation between the three process mining techniques and event logs [14].

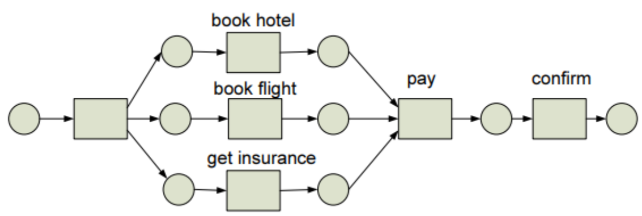


Figure 2: Example of a business process visualized with a Petri net [19].

business process twice, Figure 2 shows a visualization using a Petri net and Figure 3 a visualization using BPMN [19].

2.1.2 PM4PY

For our PoC, a process mining tool has to be used for generating process models and executing analyses. The selected tool for the PoC is PM4PY (Process Mining for Python), we selected PM4PY because it is open-source, regularly updated, and has built-in functionality to send prompts to an LLM. PM4PY is a process mining tool based on Python, this tool aims to bridge the gap between process management and data science [8]. Currently, it has over one million downloads globally [7], the tool is designed for use in both business and academia. PM4PY supports multiple algorithms for process discovery (e.g. Alpha miner, ILP miner, etc.) and it supports multiple process model formats, including Petri nets and BPMN [20].

2.2 Large Language Models

Large Language Models (LLMs) are the main component of the proposed approach to generate domain knowledge. In recent years, we have seen rapid development within the field of Artificial Intelligence (AI), especially in generative AI. LLMs, such as Generative Pre-trained Transformer (GPT) models, are the latest developments in this field [21]. These models are trained on large corpora of text data and can produce human-like textual outputs instantly.

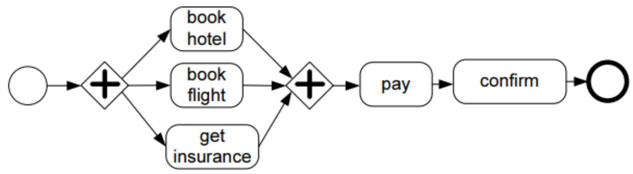


Figure 3: Example of a business process visualized with BPMN [19].

The training of these models happens with large training data sets, that consist largely of publicly available data. However, to enable an LLM to answer more context-specific questions, context can be presented to the model in the prompt, or the model can be fine-tuned [22]. Fine-tuning can be done by retraining the model on new context-specific data, which typically requires large amounts of training data [23]. An alternative approach is to use prompts to feed the model with the new contextual data. The set of techniques that are used for creating prompts is called 'prompt engineering'.

2.2.1 Prompt engineering

Prompts are the go-to way for communicating with LLMs, people use prompts to receive the desired output from the LLM. White et al. defined prompts as: "Prompts are instructions given to an LLM to enforce rules, automate processes, and ensure specific qualities (and quantities) of the generated output." [24]. Using the right prompting strategy can have a big impact on the quality of the generated input [25]. Which prompts work the best depends on the user, the intended goal, and the context. In this section, we present the prompt engineering techniques that are used in our PoC and in Section 6 we describe how these techniques were implemented in our PoC:

- **In-Context Learning (ICL)**: A technique where task demonstrations are presented to the model and based on these demonstrations the model should be able to solve a new task. The effects of this technique can be similar to fine-tuning the model, depending on the choice of the demonstrations. Dong et al., describe ICL as: "First, ICL requires a few examples to form a demonstration context. These examples are usually written in natural language templates. Then, ICL concatenates a query question and a piece of demonstration context together to form a prompt, which is then fed into the language model for prediction." [26]. Figure 4 shows an illustration of ICL.

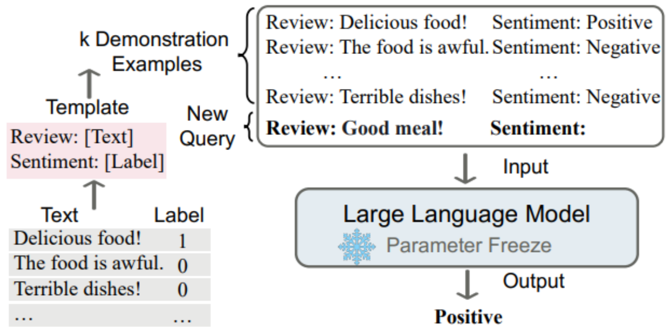


Figure 4: Example of using ICL [26]

- **Chain of Thought (CoT):** Involves breaking down a bigger problem into smaller reasoning steps [27]. CoT can result in better performance of the model, Figure 5 shows the comparison between a traditional prompt and a CoT prompt. CoT and ICL can be combined if multiple demonstrations for each intermediate reasoning step are presented to the model. The main advantage of using CoT is the potential performance improvement but the disadvantage is the required manual crafting of the prompts ('manual CoT').

- **Retrieval Augmented Generation (RAG):** A technique that can be used to feed data from an external source into the LLM [28]. This allows for the usage of real-time data in the LLM and can potentially increase the quality of the output. By using RAG the necessity to retrain the LLM on new data can be avoided, which can save time and costs [29]. For instance, RAG can be used to create a chatbot that knows the specific catalog of products of a company or that can tell you which flights are delayed departing from New York. The original user prompt is combined with retrieved additional data (from the external source) and that new prompt is processed by the LLM, as shown in Figure 6.

2.2.2 Agents

AI agents are entities that can make decisions and take actions, based on their perception of their environment [31]. The development of LLMs led to an increasing interest in LLM-based agents, as these models have the potential to allow for the creation of more powerful agents. LLM-based agents can interact via natural language with a human user, making the agent easier to use and more explainable [32], and they can leverage generative capabilities to create plans and understand tool capabilities.

Most agents contain a profiling, memory, planning, and action component. The profiling component determines the characteristics of the agent and the memory component allows the agent to remember the original task and output of previous actions. With the planning component the agent can generate a plan to solve the original task and the action component allows the agent to execute tools. Tools enable the agent to perform functions (e.g. book a hotel for you) instead of only receiving and sending natural language.

A technique that can be used to create an LLM-based agent is Reasoning and Acting (ReAct). ReAct is an approach that combines two approaches, reasoning (CoT) and acting, to achieve synergies [33]. This distinction between reasoning and acting tries to mimic the human cognitive

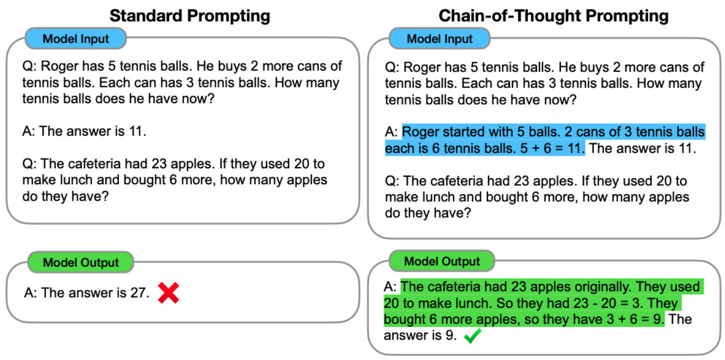


Figure 5: Example of a traditional and CoT prompt [27]

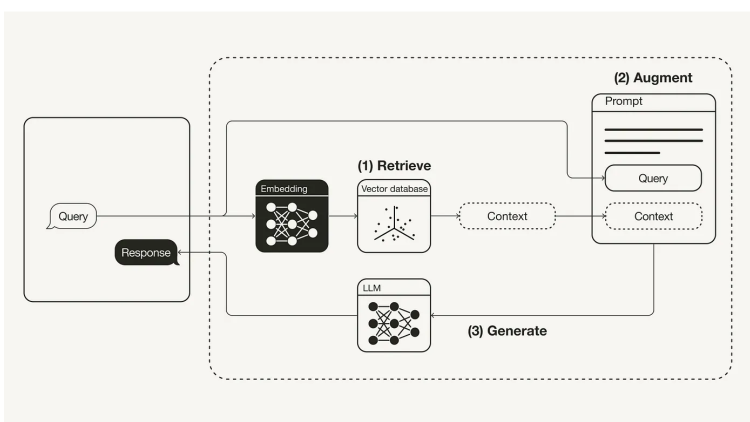


Figure 6: Concept of RAG [30]

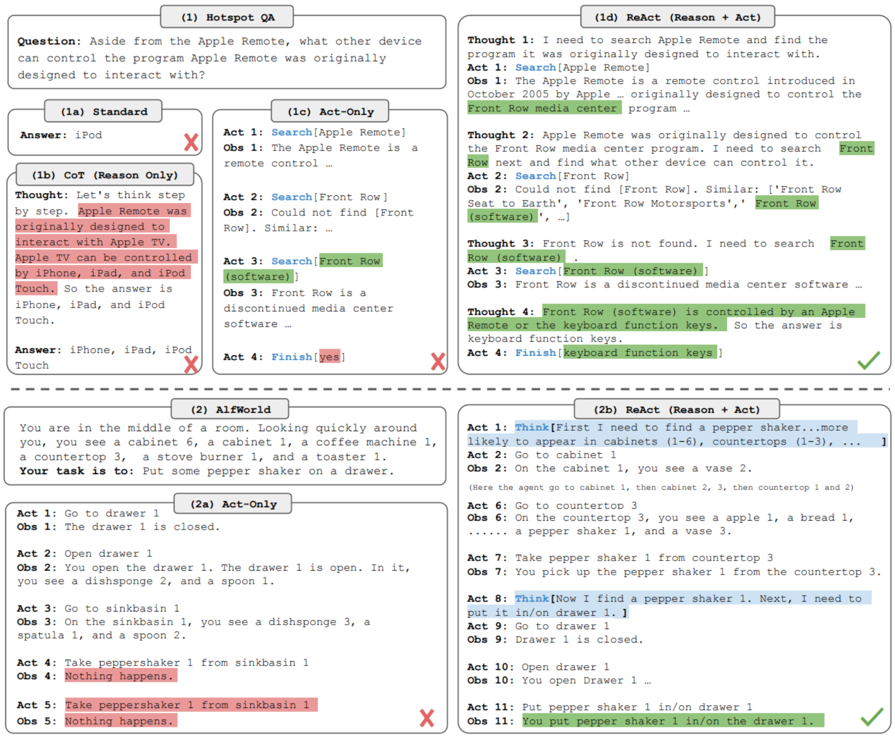


Figure 7: Comparison of ReAct to other prompt engineering techniques [33]

process. The concept is that reasoning helps the model to track and update a plan of action and handle exceptions while acting makes it possible for the model to interact with external sources such as tools or knowledge bases. This can allow the model to successfully interpret and execute more complex tasks that can also involve interactions with external sources or tools.

Figure 7 shows an example of ReAct prompting in practice, compared to other prompting techniques. The model first generates a 'thought' about the current state ('reason') and what action it should execute next ('act'). Then it executes this action, which often entails executing a tool that was provided to the model (e.g. an external function). Then the model observes the outcome of this act ('observation'), the returned value of the executed tool. It keeps going through these steps until the model decides that it has completed the initial task, then the final answer is returned to the user.

Reasoning WithOut Observation (ReWOO) is an LLM-based agent framework proposed by Xu et al. as an improved version of ReAct [34]. While ReAct allows an agent to solve more complex tasks there are some drawbacks. For instance, ReAct makes calls to the LLM at each step of the process, creating more computational complexity and increased token usage.

ReWOO decouples the reasoning and acting parts of the agent. According to experiments reported by Xu et al. ReWOO achieved a 5x token efficiency and 4% accuracy improvement, compared to ReAct. A ReWOO agent has the following components:

1. **Planner:** Creates a blueprint to reach a final solution, consisting of 'Plan' steps and '#E' (evidence) outcomes. The evidence component allows the worker to use the outcomes from past actions, to tackle multi-step tasks.

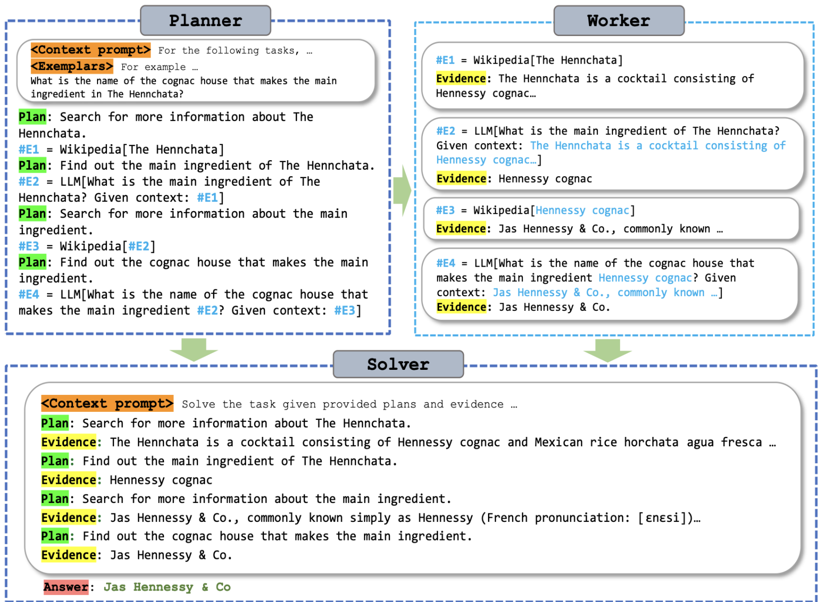


Figure 8: Example of the input and output of the planner, worker, and solver for ReWOO [34]

2. **Worker:** The instance that makes the iterative calls to the tools of the agent, based on the generated plan of the Planner instance.

3. **Solver:** This instance processes the plan and the outcomes of the worker to generate a final solution for the initial task.

Figure 8 shows examples of the ReWOO approach, with the mentioned components of the agent. An advantage of using the ReWOO approach is the lower token usage and amount of LLM calls, compared to ReAct. Figure 9 shows the two approaches and their interactions with an LLM. ReAct is a way of 'Reasoning with Observation', the intermediate observations are included in LLM calls. ReWOO only utilizes an LLM in the planner and solver component, instead of communicating with the LLM during the intermediate 'Worker' steps.

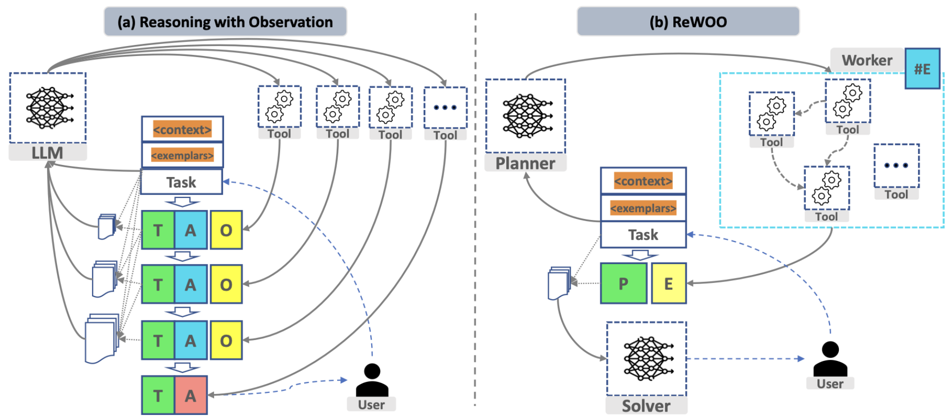


Figure 9: Overview of Reasoning with Observation and ReWOO approaches [34]

2.3 Technology acceptance

For our PoC, it is important to look at when users accept or reject the output given by a system. The unified theory of acceptance and use of technology (UTAUT) is a model that can be used to assess the technology acceptance of a system [35], shown in Figure 10. This model allowed us to structure the topic of technology acceptance when we discussed this during our expert interviews (Section 5). The UTAUT has four main determinants of intention and usage and four moderators of relationships [36]. The UTAUT tries to explain the behavioral intention and the usage behavior that follows from it. The four main determinants are:

1. **Performance expectancy:** The extent to which an individual expects that using the system will have a positive impact on job performance.

2. **Effort expectancy:** The extent to which an individual associates ease of use with the system.

3. **Social influence:** The extent to which an individual believes that important others think that he/she should use the specific system.

4. **Facilitating conditions:** The extent to which an individual thinks that there is support for using the system, both from a technical and organizational perspective.

The four moderators of the relations between the determinants and the behavioral intention and usage behavior are gender, age, experience, and voluntariness of use.

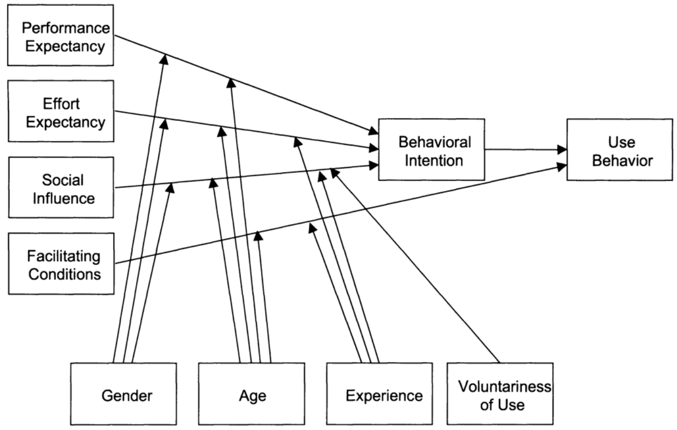


Figure 10: The UTAUT model [36]

2.3.1 Process mining adoption

In recent years, process mining has been receiving more attention, this can for instance be observed in the growth of commercial process mining tools and companies (e.g. Celonis). The global market for process mining is expected to grow from over $1 billion in 2023 to over $27 billion in 2030 [37]. Syed et al researched the way process mining is adopted within an organization. They examined this using the theory of technology discontinuity, this theory aims to explain how changes in technology can change the organization [38] [39]. From the study, four enablers of process mining adoption and seven challenges for process mining adoption were identified. Process mining should (amongst other things) gain the trust of the users, deliver actionable insights, and provide a perceived benefit to allow for a successful implementation. And users should receive the proper training to use these tools and interpret the results they produce. The challenges that were identified included the misinterpretation of the results by users, users' resistance towards using process mining due to the perceived limitations of the systems, process complexities, and user challenges related to resistance to new technologies and their perceived lack of ability to use such a system.

Besides moving to a more continuous adoption and broader adoption of process mining, process mining adoption can also be viewed as the degree to which multiple process mining techniques are used. Reinkemeyer argues that the future expansion of process mining should not just focus on lateral expansion (more processes and discoveries) but on vertical expansion (adding more depth and business value in places where process mining has already made a positive impact) [40]. To achieve this process mining implementations should have a clear purpose, engage the right people, and deliver strong technical performance.

3 Related work

In this section, we present the related work on using LLMs in process mining and supplementing domain knowledge into process mining tools.

3.1 Using LLMs in process mining

We make a distinction between two types of studies that are focused on using LLMs within process mining. The first one is studies that outline the potential opportunities and challenges of using LLMs in process mining, functioning as a research agenda. The second one is studies for which a prototype was developed or practical experiments were conducted with LLMs.

We now discuss the first type of studies. Klievtsova et al. studied the potential of AI-driven chatbots in five phases of the process life-cycle (gather information, process modeling, assure model quality, select redesign method, and apply redesign method) [41]. They conclude that chatbots can already have a considerable business impact in process modeling, the information gathering and process modeling phase. However, future research should focus on integrating the language capabilities of AI-driven chatbots into more specialized process mining tools. Busch et al. provided an overview of the potentials and challenges of using prompt engineering techniques in BPM, compared to using fine-tuning [42]. They concluded that using prompt engineering can partially fix the issue of a lack of labeled training data, which is required for fine-tuning the model. Furthermore, prompt engineering can offer increased explainability and improved computational efficiency.

Now we discuss the second type of studies, for which LLMs were used in a practical experiment or prototype. For the PM4PY tool, functionality has been developed to connect the tool with an LLM [7]. This functionality allows the LLM to interpret the process models via textual abstractions of those process models [43]. The LLM can then answer questions of the user about the process model. It does not (automatically) utilize more advanced prompt engineering techniques, instead it uses a 'zero-shot' prompting approach [44]. For instance, a user can ask the LLM descriptive questions about the process model. This solution is now just focused on providing information about the process model and data to the user via the LLM.

Jessen et al. developed an approach for using LLMs in process mining, aiming to make process mining tools more accessible [45]. Their approach also focuses on utilizing LLMs to create a conversational agent that can answer questions about the process model or event log. Their conclusion was that LLMs can function as the basis for such a system but that improvements can be made in the future. Kourani et al. proposed a system called 'ProMoAI' [46], where the LLMs are used to generate process models based on textual descriptions of the process. While this type of research also focuses on utilizing LLMs within process mining, the task they are used for differs from our approach. Another approach was taken by Grohs et al. who also studied LLM applications within process mining [47]. They focused on utilizing an LLM for mining the process from textual descriptions, so during the process discovery phase. They concluded that LLMs were able to execute these tasks. Rebmann et al explored the potential of using LLMs to execute semantic-aware process mining tasks, e.g. anomaly detection or next activity prediction [48]. They concluded that LLMs often fail to successfully complete these tasks without fine-tuning. However, if the LLMs are fine-tuned the performance improves.

If we look at initiatives from the business world, we see that process mining software vendors are moving towards integrating LLMs in their process mining tools. Some examples of large process mining software vendors and their approach:

- **SAP:** Developing a Large Process Model (LPM), that combines the capabilities of an LLM with SAP's experience in business processes [49]. This will be part of SAP Signavio, the process mining offering from SAP.

- **Pegasystems:** In May 2023, Pegasystems presented Pega APIs which make it possible to connect their interface to generative AI models [10].

- **Celonis:** Has demonstrated generative AI capabilities for their process mining, during the Celonis 10-city World Tour [9]. All of the demonstrated functionality was part of a demo and is not available for commercial use yet. The capabilities that were demonstrated include parsing data and visualization requests using Natural Language Processing (NLP) requests.

From the first type of studies, we observe that the authors indicate that AI-driven chatbots and using prompt engineering techniques (in combination with LLMs) can have a positive impact on process mining. From the second type of studies, we see LLMs being applied for different tasks within process mining with varying results. Current efforts are mainly focused on using LLMs in process mining tools for generating process models, providing descriptive knowledge, or providing knowledge about a specific tool. Our approach proposes to use LLMs to provide another type of domain knowledge into process mining tools. Specific domain knowledge about the domain and business of the organization, e.g. based on characteristics of the sector or the specific organization. The provided domain knowledge allows a user to interpret the process model or make process improvements, making process mining tools more user-friendly, time-efficient, and accessible.

3.2 Supplementing domain knowledge

Eichele et al. studied the potential of adding domain knowledge to event logs, using Web Ontology Language (OWL) ontologies [50]. With this approach, they can map the domain knowledge to specific cases and activities. This angle differs from our approach as this is focused on the manner of supplementing domain knowledge and not on the automatic generation of domain knowledge, that we are focusing on. Therefore, these two different angles can potentially benefit from each other, our automatically generated domain knowledge could be supplemented into a process mining tool using their ontology.

4 Methodology

In this section, we present the methodologies that we have used for this research. We present our hypotheses and research objectives first. Then we present the main methodology, the methodology for the expert interviews, the methodology for the PoC, and the methodology for the evaluation of the PoC.

As mentioned in the introduction, our research question is: 'How can LLM-based agents provide domain knowledge to help users of process mining tools understand and improve processes?'. To help answer this research question, we developed two hypotheses:

1. *H_1: An LLM-based system can provide domain knowledge about the context of a process to the user of a process mining tool.*

2. *H_2: An LLM-based system can partially take over the role of a human domain expert in the deployment of process mining.*

The research objectives that we identified are listed below, these objectives help with trying to answer the main research question. The research objectives for this thesis are the following:

1. *RO_1: Develop a framework that describes how domain knowledge can be included in a process mining tool.*

2. *RO_2: Explore techniques that allow LLMs to produce useful output based on domain knowledge.*

3. *RO_3: Develop an LLM-based Proof-of-Concept (PoC) that can generate and supplement domain knowledge into a process mining tool.*

4.1 Research method

The main methodology of this study is research by design, based on the framework proposed by Peffers et al. [11]. This framework consists of a 6-step approach to execute research by design specifically for information system research. Below we list the steps of the framework and how we used them in our research:

1. Problem identification and motivation: The problem was identified based on the studied literature (Section 2) and validated with expert interviews (Section 4.2 & 5).

2. Definition of the solution's objectives: The research objectives are presented in Section 4 and are based on the mentioned literature study (Section 2) and the expert interviews (Section 4.2 & 5.

3. Design and development: The design and development of the system are discussed in Section 6, where we present the conceptual framework and the PoC that we developed.

4. Demonstration: The demonstration of the developed system is presented in Section 6. In this section, the technical architecture and components are presented.

5. Evaluation: The evaluation of the system is done through an ablation study and qualitative analysis of the output of the PoC. We outline our approach for the evaluation in Section 4.4 and in Section 7 we present the results of the evaluation.

6. Communication: The communication of this research is discussed in Section 9 and in Section 8 we discuss potential limitations of our work and potential directions for future research.

4.2 Expert interviews

Based on the studied literature, we identified a lack of domain knowledge as an issue in the current state of process mining. To validate this observation, we conducted a series of seven expert interviews. These interviews were semi-structured, meaning that a list of pre-defined questions was prepared before the interview but at the same time there was room for other follow-up points or discussions. The pre-defined questions that we used for the expert interviews are included in Appendix A. Since these interviews were explorative, we did not make use of coding to analyze the transcripts. The interviews aimed to get input from process mining practitioners on the current state of process mining tools and to see if and how they thought that LLMs (generative AI) could have a positive impact on that. We used the input from the interviewees as the basis for the framework and PoC that we developed. The list of questions was divided into five topics:

1. **Background:** General questions about the education and professional experience of the interviewee.

2. **Process mining & domain knowledge:** Questions about the interviewee's perception of the current state of process mining and potential obstacles or issues that the interviewees had observed in the current state of process mining. This section also included questions about domain knowledge and the role it can play within process mining. This section aimed to get input on the state of process mining and to confirm or deny the mentioned issue of a lack of domain knowledge (and potentially identify other issues) in the current state of process mining.

3. **Generative AI:** This section included questions to get insights into how the interviewees had experienced their past interactions with generative AI models. And to see if they see potential for generative AI in process mining and how they think the current generative AI models can improve. This section aims to get input on the perception of generative AI from the interviewees and to get input on the potential use cases for generative AI within process mining.

4. **PoC & technology acceptance:** Two questions to assess which factors influence the decision of the interviewees whether they accept or reject the output of generative AI models and new technologies in general. This section aims to identify factors that influence the technology acceptance of the interviewees, as these factors can be important for the future acceptance of our PoC and its output.

5. **Closing:** Room for the interviewees to voice any other ideas or concerns they had about one of the topics or other things.

For the expert interviews, we selected a diverse set of process mining practitioners to get a complete view of the discussed topics. In this context, we mean by diverse that the interviewees have different roles within the world of process mining. For instance, a process mining developer can give insights into how new solutions can be developed and what is already possible. Whereas, a

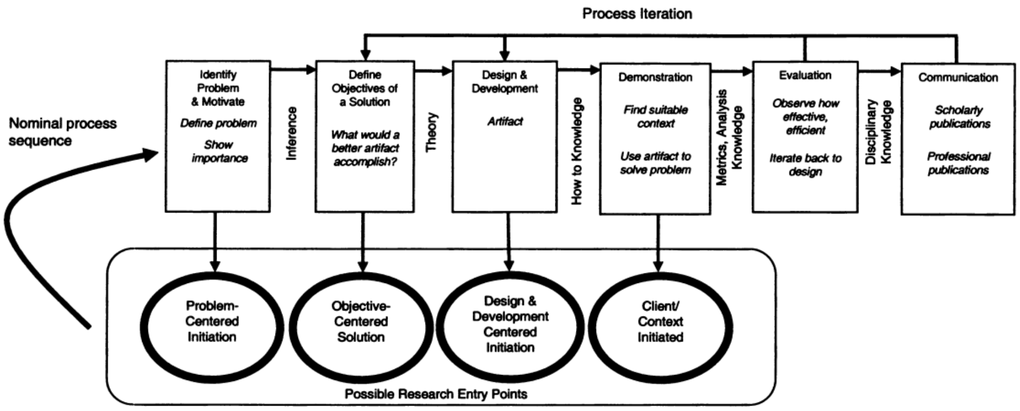


Figure 11: The DSRM Process model [11]

process mining end user can indicate where the current process mining tools are falling short for people without a full technical understanding of these tools. People who hold managing positions related to process mining can give input about the strategic use and potential of process mining. The interviewees included a diverse set of process mining practitioners, Table 1 (Section 5) gives an overview of all the interviewees.

The answers that the interviewees gave are their opinions and observations and are therefore subjective. These answers to the interview questions can be prone to personal biases, to (partially) mitigate this risk we conducted multiple expert interviews with experts from different organizations, countries, and job levels. A higher number of interviewees with different backgrounds lowers the risk of these biases occurring in the results and increases the evidentiary value of these results [51].

4.3 PoC

The methodology for the design and development of the PoC is 'research by design' and we used the framework proposed by Peffers et al. [11], shown in Figure 11. The first step of this framework is the 'problem identification and motivation' step, for this step we performed a literature study and conducted expert interviews. The literature study and expert interviews were also used for the second step of the framework ('define objectives for a solution'), based on those we developed the presented research objectives (Section 4). Section 4.2 describes the specific methodology used for these interviews and Section 5 describes the results that were gathered from these interviews.

The third step in the framework is the 'design and development' step, where the system or PoC should be designed and developed. Based on the results of the first and second steps of the framework we present the design and the developed PoC in Section 6, aligning with the 'demonstration' step. In Section 4.4, we present how we evaluated the PoC, aligning with the 'evaluation' step of the framework. The last step is 'communication', we cover this in Section 9, where we deliver our conclusions.

4.4 Evaluation

In this section, we discuss the methods we used to evaluate the PoC. We conducted two experiments, an ablation experiment and a qualitative analysis of the output of our PoC across different executions of the PoC.

4.4.1 Ablation experiment

An ablation experiment, in the context of AI-based systems, involves removing or changing components of the system's architecture and observing its performance to gain a better understanding [52]. We tested three different ablations to compare their outputs. The PoC that we developed is a ReWOO-based agent that generates domain knowledge. The process mining capabilities come from the PM4PY module and the domain knowledge capabilities from the GPT Researcher module [8] [53]. To test whether or not our PoC offers additional value compared to these two components of the PoC, we tested the following three ablations:

1. PM4PY LLM query functionality [7]
2. GPT Researcher [53]
3. Our PoC (Section 6.2)

We selected these three systems because PM4PY represents the process mining capabilities and GPT Researcher the domain knowledge generation capabilities of our PoC. We compared these two to our PoC to see if the construction of our PoC offers additional value compared to these separate components of our PoC.

4.4.2 Qualitative analysis

With the qualitative analysis, we evaluate if the PoC performs differently across various problem identification types and use cases. This is important for determining whether or not the PoC offers additional value in all areas or specific ones. Furthermore, we also use different process types (different event logs) and different techniques (e.g. DFG) to see if this influences the results.

We let our PoC generate a report and then we analyzed how much of the content of the report contains information that is specific to the process component that the system has analyzed. We used this approach since LLMs and LLM-based systems tend to generate generic knowledge and therefore a challenge is to let such a system generate specific information. An example of generic information could be the system suggesting to remove manual processing (which can be applied to many processes and activities). An example of more specific information could be the impact of return policies on the carbon footprint of an organization (e.g. for the 'Return Goods' step in a Purchase-to-Pay process).

5 Results from expert interviews

In this section, we discuss the results of the conducted expert interviews that were used as a basis for the PoC. We conducted semi-structured expert interviews to get input from process mining practitioners, the list of questions that was used can be found in Appendix A. These questions were divided into five topics, in the next sections the answers to the questions of the main three topics are discussed. The first and last sections (background and closing) are not discussed because these did not provide results that are relevant to answering the research question of this study. Table 1 shows the interviewees of the conducted expert interviews.

5.1 Process mining & domain knowledge

When the interviewees were asked to name some barriers or obstacles in the current state of process mining, multiple ones were identified:

- **Unknown technique (Subject B):** Process mining as a technique and its capabilities are still relatively unknown, which limits the adoption of process mining.
- **Complexity (subject B, C, and E):** The complexity of process mining can be a potential barrier to the adoption of process mining. People often find process mining intimidating if they do not know the technique.
- **Business value (Subject A):** The current process mining tools sometimes fall short in solving a business problem for an organization. The tools are often developed in academia and should be focused more on adding value to organizations.
- **Unavailability of data (subject C, D, E, and F):** Organizations are not willing to share all the data or the data of the systems is not ready to be used for process mining.
- **Process unawareness (subject D):** People within organizations are often not aware of the complexity of their processes and because of that think that process mining is not necessary in their organization. Subject D described this as:

  "Many clients believe that their process is not as complex as we try to tell them that it is. They do believe that they have a 100% understanding of what their process is. So they do not see the value add of process mining"

  However, those processes turned out to be more complex than the business owners thought. They for instance thought that their process had just a couple of variants but process mining proved that there were thousands of them.

All interviewees agreed on the fact that domain knowledge about the context of the process and the organization is an important asset in process mining, although their arguments differed. For instance, subject A argued that theoretically domain knowledge should not be required, if the tools were good enough:

Interviewee	Job and organization	Process mining experience	Generative AI experience
Subject A	Data Scientist, Pegasystems	Mostly during academic study and research, wrote his master thesis about process mining	About two years experience, both professionally and personally, as a user and developer
Subject B	Process Intelligence Consultant, Apolix	one year of professional experience (mainly focused on implementation work)	Using ChatGPT since its release for both personal and professional usage
Subject C	Fellow Data Scientist, Pegasystems	Around five years of professional experience (mainly research & development)	About two years experience, both professionally and personally, as a user and developer
Subject D	Consulting Manager, Pegasystems	Around one year of professional experience, focused on managing process mining teams and projects	About two years experience, both professionally and personally, focused on potential use cases
Subject E	Senior Consultant, Deloitte	About two and a half years of professional experience, completed around 12 client engagements regarding process mining	Using ChatGPT since its release as a user and researching the potential of LLMs for client engagements (as a developer)
Subject F	Partner & CTO, Deloitte	Around 15 years of professional experience with process mining	About two years experience with generative AI, mainly focusing on identifying use cases and developing strategies

Table 1: Overview of the interviewees

"I think if you were to look at it as a purist, no. Because I think that if those analytical tools are good enough, they can also do it without domain knowledge. But the realistic answer is that those analytical tools are not good enough, so a lot of interpretation is often needed to achieve a better result."

According to the interviewees, the required domain knowledge is both knowledge about process mining itself and knowledge about the context of the process. All of the interviewees agreed on the point that domain knowledge about the context of the process is often required to understand the process and to be able to perform process enhancement, or as subject D put it:

"But what I think that is still very important, is that we have people on these projects that understand what the substance of these processes actually is. Because the lead time of a step in one process can be a big problem, that you have to fix, that same lead time can be very normal in another process."

Subject F mentioned that in practice domain knowledge can often explain certain things in a process that from a process mining perspective seem inefficient or special:

"I have hundreds of examples, where when you are conducting an analysis and you think: now I have really found something, if I show this to the client they can save millions. But then you show it to them and it turns out that this is part of their business model."

For instance, a consumer goods company that has zero value deliveries when they give away free stuff. From a process mining perspective, one could argue that this is inefficient and that it should be fixed. However, that can be part of the marketing and promotions strategy of the company. Subject D also mentioned that if the process mining tool could provide a little information about the domain, that could help the implementation of process mining tools:

"But if we now staff consultants on a project where they quickly have to deliver input based on the data, they will have to heavily rely on domain knowledge of people at the client. And I can imagine that if the tool has a little more understanding of those specific domains, that can definitely help."

When asked if and how this domain knowledge is already present in current process mining tools, different answers were given. Subject A argued that often the domain knowledge is present within the data of the event logs but that explicit domain knowledge currently is not present in process mining tools. Subject B explained that in some process mining solutions, there currently is domain knowledge present about process mining itself. He gave a short walk-through of the functionality that he mentioned, which was an 'about this' page within a process mining solution. This page gave additional information about some of the functions of the specific tool, so this would fall under 'process mining domain knowledge' (Figure 12). No interviewee indicated that 'ecosystem domain knowledge' (Figure 12) is currently present in process mining tools. Subject D described this as:

"As far as I have seen, those tools are pretty agnostic, right? Currently, it is not the case that the tool is able to tell you, this is a financial process and that it then comes with a different analysis output or that it tells you in that case you should focus on these aspects."

Subject A and subject E also mentioned that there is a risk of having too much domain knowledge, which can cause a person or organization to get stuck in a certain situation. Where processes have been the same for a long time and people think that they have a lot of reasons to leave these processes unchanged. This risk does not apply to our approach, as we generate new domain knowledge and provide that to the user. The knowledge that the interviewees referred to is knowledge based on experiences that people have built over many years of experience. Some interviewees (B, C, E, and F) also mentioned that in general, the knowledge about process mining techniques at most organizations is quite low.

5.2 Generative AI

All interviewees indicated that they had used generative AI solutions, most often ChatGPT of OpenAI was named [54]. The general experiences with LLMs were positive according to the interviewees. Prompting techniques were mentioned as a big influence on the quality of the output of the LLMs, e.g. using ICL (Section 2.2.1). The risk of encountering hallucinations when using a LLM was mentioned as a drawback of using LLMs, by subject E.

The interviewees proposed multiple use cases for LLMs within process mining:

- **Analysis chatbot (subject A, B, C, and F):** A chatbot that acts as a layer on top of the process mining tool, the user can ask the chatbot questions about the process model ('Descriptive tasks', Figure 13).

- **Code writer (subject B):** Some implementations of process mining tools require coding in specific languages, such as Process Query Language (PQL) for Celonis [55]. This is a vendor-specific language and most users will be unfamiliar with PQL, therefore using an LLM to write PQL code could be beneficial for users. This refers to knowledge about a specific tool, so this is 'process mining knowledge' (Figure 12).

- **Process enhancement (subject C):** An LLM-based chatbot could prove useful in enhancing a process model when a user does not know how to fix an issue:

  > *"Maybe generative AI can give you advice on what to do next. Maybe you found some scenario that you're stuck in, okay, I've found this problem. Maybe generative AI could take you to the next level and the next step."*

  This idea is incorporated into our conceptual framework in the 'improvements' component (Figure 13).

- **RAG-driven generative AI model (subject D):** Using RAG to supplement specific knowledge about the processes, to help understand the processes and the output better. As mentioned in Section 6, we applied RAG in our PoC by using a ReWOO-based agent with tools that can insert external knowledge.

- **Script generator (subject E):** Letting a LLM generate scripts that transform the data of a company into usable event logs for process mining.

Some interviewees expressed that they were quite satisfied with the current performance of the LLMs that they used. Others said that especially the quality of output of the LLMs should improve in the future, to maximize the value that they offer.

5.3 Prototype/ technology acceptance

When asked whether or not the interviewees always accept the output of LLMs, all said that they do not always accept the output. Subject B mentioned that most people tended to accept most of the output of LLMs when they were released initially. This observation aligns with the 'social influence' determinant of the UTAUT model (Section 2.3). In the beginning, people started using generative AI solutions and that created a hype and more and more people started using it. Currently, people are realizing more and more that LLMs can give incorrect output. Another issue that was mentioned by subject A, was the often lacking ability of LLMs to identify questions to which they do not know the answer. According to the interviewee, when the model does not know the answer, it starts to hallucinate. Implementing guardrails, e.g. using prompting techniques, was mentioned by subject C as a measure to increase the quality of the output.

Multiple interviewees (subject A, C, E, and F) indicated that the user experience has a big influence on the general adoption of new technologies. They argue that the underlying technology is less relevant, users are mainly looking for additional value, that outweighs the costs of using the technology. And that new technologies need to deliver correct and trustworthy output, immediately when they are released. When the quality is insufficient at its release, it will be difficult for a new technology to convince people of the added value at a later stage. As subject C described it:

"I think when you're introducing it to a new client or you're trying to build something for somebody, you need to have something that is good enough to where they're going to say oh wow, I really see the value in this."

This corresponds with the 'performance expectancy' and 'effort expectancy' determinants of the UTAUT model that was discussed in Section 2.3 [36]. Subject B said that age can also influence people's technology acceptance, younger people would be more willing to accept new technology compared to older people. Age was also a moderating factor in the UTAUT model for technology acceptance.

5.4 Summary of findings

The conducted interviews gave a diverse set of perspectives of process mining practitioners. The key insights that were gathered during the interviews and that are most relevant to this study are:

1. **Process mining obstacles:** Currently, multiple things limit the overall usage and adoption of process mining. From the mentioned obstacles complexity and business value are factors that might (partially) be solved by this research. The proposed framework and PoC aim to make the usage of process mining tools more accessible by using input and output in the form of natural language, which can be easier to interpret than large process models. Adding the domain knowledge into the tool can make the results more usable and actionable, increasing the business value. A process model on its own can be hard to interpret without the relevant domain knowledge.

2. **Domain knowledge & process mining:** Domain knowledge about the context of the process was identified as an important factor by all interviewees. No interviewee knew an example of a process mining tool, that already explicitly incorporates ecosystem domain knowledge in the tool. That finding aligns with the observation made in the literature study

(Section 2), that ecosystem domain knowledge is important in process mining but it is not yet represented in current process mining tools. Therefore, there is a knowledge gap, which this study tries to (partially) fill.

3. **Generative AI experience:** The majority of the interviewees expressed that they were (to some extent) happy with their interactions with LLMs. Prompting techniques were mentioned multiple times as a factor that can have a big influence on the quality of the output. Therefore, we explored and used different prompting techniques in the developed PoC.

4. **Process mining & generative AI:** Multiple potential use cases for generative AI in process mining were mentioned by the interviewees. One approach that was proposed by an interviewee, was using a RAG-driven model for providing domain knowledge. We incorporated some of the use cases for LLMs in process mining that the interviewees proposed into our framework (Figure 13). We described which ones we incorporated in Section 5.2.

5. **Technology acceptance:** The main factors that influence the acceptance of technology (according to the interviewees) are the user experience and the added value. Users should immediately see the benefit of using a new technology, compared to the old situation. This aligns with the 'performance expectancy' determinant of the UTAUT model and we took this into account when developing our framework and PoC. We incorporated this factor by focusing on a system that could generate a complete set of domain knowledge that is specific to a component of the process.

We used these findings for the development of the framework that we proposed and the PoC that we developed (Section 6).

6 System design

In this section, we present our conceptual framework and PoC of an LLM-based system for supplementing domain knowledge into process mining tools. The framework and PoC are based on the findings from the studied literature and the findings from the expert interviews. Section 6.1 presents our conceptual vision of such a system and Section 6.2 presents the PoC that we developed.

6.1 Conceptual framework

We designed a conceptual framework that shows where and how domain knowledge, generated by an LLM, can be incorporated into a process mining tool to solve the observed scarcity of domain knowledge in process mining. The value of domain knowledge was validated through the literature (Section 1) and through the conducted expert interviews (Section 5). In this framework, we incorporate different types of domain knowledge and general process mining techniques.

6.1.1 Types of domain knowledge

Domain knowledge is the information that will be presented to the users of the prototype and therefore clearly defining domain knowledge is important for this research. We distinguish three different types of domain knowledge (Figure 12):

1. **Process mining knowledge:** Domain knowledge can be defined as knowledge about the domain of process mining itself, which can be an inaccessible field for non-experts [3]. For instance, knowledge about the usage of a certain process mining tool or process mining techniques, e.g. conformance checking or process discovery.

2. **Process knowledge:** Another way to define domain knowledge is as knowledge about processes and process optimization, regardless of process mining and the used process mining tool, and to some extent independent of the specific problem domain or industry. For instance, knowledge about Lean Six Sigma, a set of techniques for process improvements and waste elimination [56].

3. **Ecosystem knowledge:** Domain knowledge can also be defined as knowledge about the domain and business of the organization, the ecosystem surrounding the process. This knowledge is important when interpreting the process model and when trying to make changes to the process. The lack of this knowledge has been identified in the current state of process mining through our literature study and expert interviews [4]. Within this type of domain knowledge, we make the distinction between external and internal ecosystem domain knowledge. Internal ecosystem domain knowledge consists of knowledge only available within an organization, this can be any type of internal documentation or other types of information sources. External ecosystem domain knowledge is publicly available knowledge, e.g. regulatory documents or sector reports published online.

6.1.2 Framework

In this section, we explain our vision of a system that can supplement domain knowledge into process mining tools, by including the types of domain knowledge and the different process mining

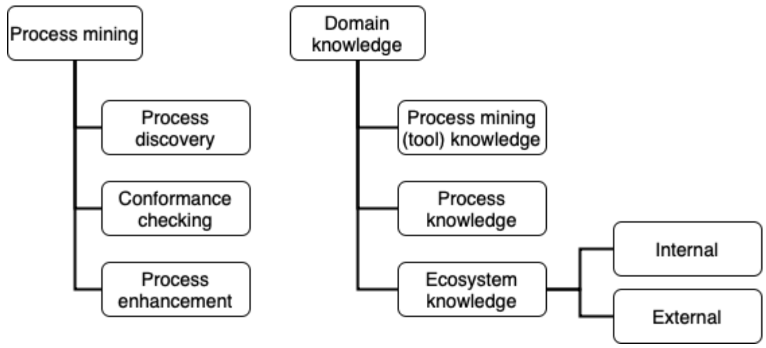


Figure 12: Process mining techniques and types of domain knowledge

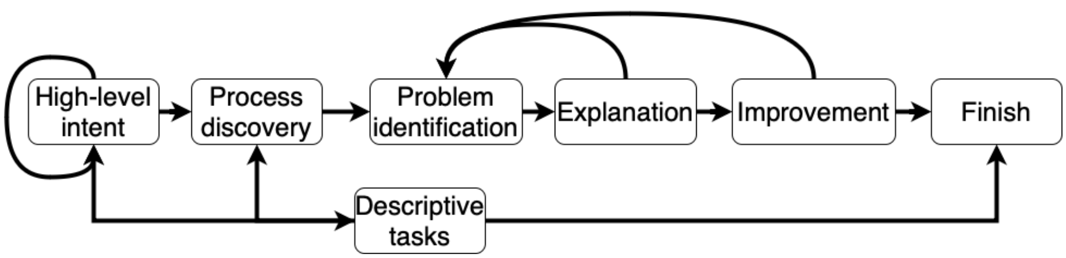


Figure 13: Conceptual framework

techniques into a conceptual framework. Figure 12 shows an overview of the types of domain knowledge and the general process mining techniques.

We propose a conceptual framework that indicates where and how LLM-generated domain knowledge, can be injected into a process mining tool to fill the observed gap in domain knowledge in the current state of process mining. Our framework is shown in Figure 13, we include the types of domain knowledge and the process mining techniques (Figure 12) within the parts of the framework.

The framework (Figure 13) consists of seven main components. The first one is the 'high-level intent'. Here, the user specifies what the main objective of the analysis is and potentially what tools should be used in this analysis. This can be done by using a single user prompt or by the user having a conversation with the system. In this conversation, the user can express the desired output and tools of the analysis, ask clarifying questions to the system, and the system can inform the user about the available options and what those options entail. Furthermore, the system can ask questions to the user to allow for the most complete picture of the high-level intent, such as asking the name of the organization or the type of process.

In the 'process discovery' component, the system will generate the process model from the event log that was provided by the user in the high-level intent phase [14]. If the user specifies a preference for a specific approach, that approach will be used, otherwise, the system will pick an approach for process discovery, based on process mining domain knowledge.

The third component of the framework is 'problem identification', the system will analyze the generated process from the previous step and identify the relevant information. What this relevant information is and how it is identified depends on the preferences of the user from the high-level intent. If the user does not specify a preference, the system will determine what type of analysis should be used, based on ecosystem domain knowledge. For instance, for spotting process inefficiencies, an analysis based on the frequency and duration of the process and its activities could be the best option. Another example of a technique that could be used in this phase is conformance checking [14].

In the fourth component, the 'explanation', the system generates relevant ecosystem domain knowledge for the user based on the results of the problem identification phase. For instance, if the problem identification phase has identified bottlenecks in the process, the explanation phase can provide explanations about why these bottlenecks might be occurring in the process. Based on the explanations that the system provides, the user can refine the problem identification, e.g., by applying a filter.

The fifth component is 'improvements', an extension of the explanation phase where the system will generate potential solutions addressing the identified problems in the process, aligning with process enhancement [14]. For instance, process improvements for removing the bottlenecks in the process could be generated in this phase. The user can refine the problem identification during this phase, e.g., by letting the system update the process with the proposed process improvements and execute problem identification on the updated process. If the user does not want to refine the problem identification and is content with the output from the system, the final result will be generated and returned to the user (sixth component, 'finish').

Besides the mentioned components, there is another one that can be used from the 'high-level intent' or 'process discovery component', the 'descriptive tasks' component. This represents the option for the user to ask descriptive questions about the event log (before process discovery) or generated process model (after process discovery). For instance, 'How many activities are between A and B in the process model?'. These are the types of questions that can be answered with process mining domain knowledge and process domain knowledge (Figure 12). The capacity to answer these types of questions with LLMs is already being studied and developed by others (e.g., [7, 45]).

With this framework, we allow the execution of all the general process mining techniques and incorporate the different types of domain knowledge. During the execution of the system, the user can still interact with the system. For instance, to change the type of analysis or apply filtering based on the generated process insights and domain knowledge. This way we do not just incorporate domain knowledge in a process mining tool but also offer the user a conversational agent that can be adaptable to the preferences of the user and answer questions. Therefore, this design should help to fill the mentioned shortage of domain knowledge in current process mining tools. The framework aligns with our first research objective (RO_1) *'Develop a framework that describes how domain knowledge can be included in a process mining tool'*.

6.2 Proof of concept

In this section, we present the developed PoC, which is based on the conceptual framework and vision of the previous section (code can be found on GitHub[1]). The goal of this PoC is to investigate the practical feasibility of a system that is based on our framework. For the PoC, we did not develop all the mentioned functionalities of the framework. The framework is a conceptual vision of a system that can execute all process mining techniques and generate all types of domain knowledge (Figure 12). For our PoC, we focused on a process mining system that can generate ecosystem domain knowledge, as the generation of this type of domain knowledge has not been studied yet.

[1] https://github.com/maxvogt12/ReWOO_agent_for_PM

Research has been done to explore the potential of LLMs to generate the other two types of domain knowledge, so we are not focusing on the generation of those types. Our PoC contains the following components of the framework (Figure 13): 'high-level intent' (as a single prompt), 'process discovery', 'problem identification', 'explanation', 'improvement', and 'finish'. The requirements for the PoC were gathered during a small set of in-depth interviews with process mining practitioners (Section 4.2 & 5).

Our PoC is a ReWOO-based agent because that allows the system to execute external functions (e.g. process mining techniques) and solve more complex tasks. With the available tools, the agent can fulfill the mentioned steps of our conceptual framework (Figure 13). From the user (in the 'high-level intent'), the system requires the following inputs:

- The file path to the event log.
- The name of the organization.
- The type of the process, e.g. order-to-cash.
- The type of analysis that should be conducted, e.g. focusing on process inefficiencies.
- The user can also express a preference for the approach that should be used for problem identification. Directly-Follows Graph (DFG), temporal profile, and variants are currently supported. If the user does not specify a preference, the system will pick one with the Technique Selector tool, see Section 6.2.1.

We chose to use an LLM-based agent and specifically a ReWOO-based agent. We made this choice because this system needs to generate multiple outputs, some of which are not natural language and these intermediate outputs are the inputs of other parts of the system. Using a ReWOO-based agent allows the smooth implementation of different process discovery and analysis approaches in the agent, from which the user can choose by simply expressing the preferences in the initial user request. It allows the system to 'reflect' during the execution, e.g. if one tool is not working sufficiently the agent can decide to retry it or try another tool. Making the system more robust in the sense that it will be able to complete the given task more often compared to a more traditional approach. Furthermore, by using ReWOO we lower the computational complexity and the token usage of the PoC (compared to ReAct). The developed PoC aligns with our second (RO_2) and third (RO_3) research objective. The second research objective states to explore techniques to let LLMs generate useful domain knowledge and the third research objective states that we should develop a LLM-based PoC to generate domain knowledge. In the next subsections, we will present the architecture and technical components of our PoC.

6.2.1 Architecture

In this section, we present the general architecture of our PoC. Figure 14 shows the 'static' architecture of the PoC. All the components of the system are displayed and one can see the three main components of a ReWOO-based agent, 'Planner', 'Worker', and 'Solver' [34]. The 'Planner' and 'Solver' instances both work based on a provided template. In this template, we incorporate ICL and CoT prompting techniques [26, 27], by providing examples with reasoning steps.

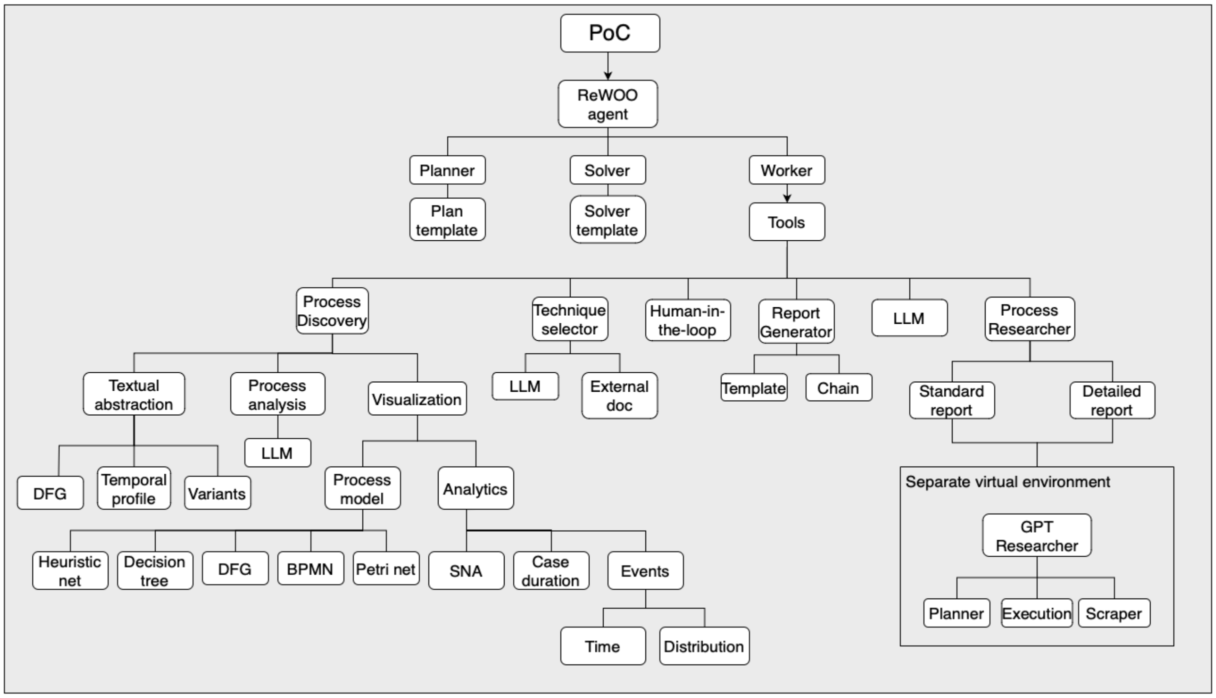


Figure 14: Architecture of the PoC

In Figure 14, the tools of the agent are shown under the 'Worker' instance, we will now discuss the purpose of each of these tools. For each tool, we will also discuss which part of our framework it represents (Figure 13). Our ReWOO-based agent has the following tools at its disposal:

1. **Process Discovery:** This tool can generate a process model from an event log and identify the relevant process components in this process model, representing the 'process discovery' and 'problem identification' steps of our framework (Figure 13). The tool uses the textual abstraction functionalities in PM4PY [43], supporting the DFG, temporal profile, and variants approach [7]. Whether the DFG, temporal profile, or variants approach is used, is determined by the 'Technique Selector' tool. After generating the textual abstraction of the process model, the tool uses an LLM to find the relevant components of a process (e.g. a specific activity in the model). These relevant components will later be used as a basis, for generating the domain knowledge.

   The Process Discovery tool can also visualize the process model for the user in multiple formats (DFG, Petri net, BPMN, heuristic net, and decision tree) and some additional statistics about the event log, e.g. the distribution of events over time. Figure 15 shows an example of a BPMN model generated by the Process Discovery tool.

2. **Technique Selector**: This tool selects the appropriate approach for the problem identification technique (DFG, temporal profile, or variants) for the analysis that the user has requested, e.g. focusing on process inefficiencies. This tool uses internal ecosystem domain knowledge in the form of a document that the tool analyses to base its decision on, simulating the use of internal documents of an organization in the execution of the system. The user can also request a specific technique that should be used by the agent, the agent will then implement

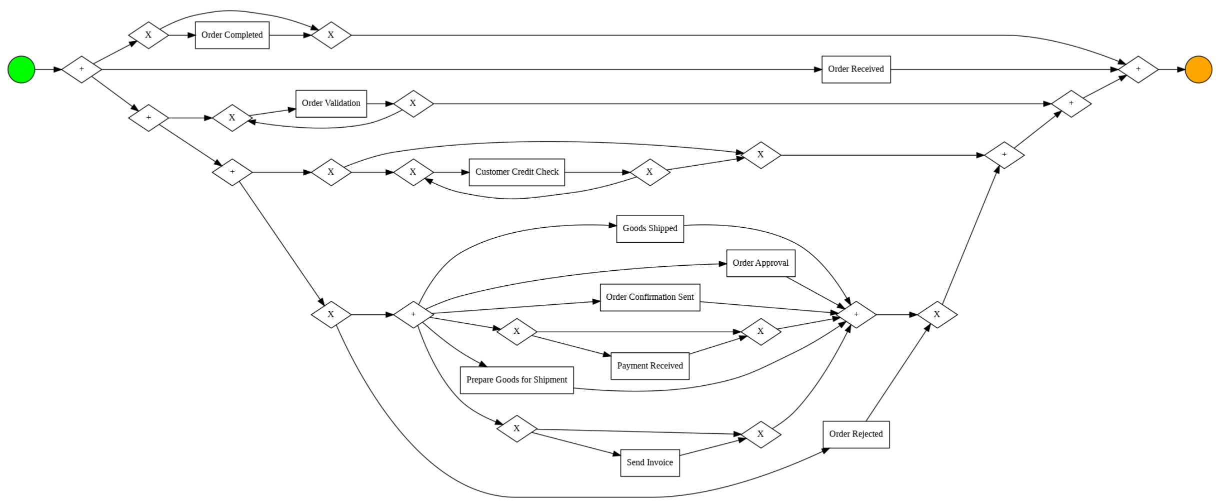


Figure 15: BPMN model generated by the Process Discovery tool (‘O2C.csv’ dataset from our GitHub page [57])

this preferred technique. If the document does not contain any useful information to base the decision on and the user does not provide a preference the agent will pick the default technique (DFG).

3. **Process Researcher:** The Process Researcher tool generates research reports about the identified process components, supplying explanations and potential improvements for the identified process issues (‘explanations’ and ‘improvements’, Figure 13). For each provided process component one report is generated, based on external ecosystem domain knowledge from online sources. The reports are generated by the GPT Researcher module, an open-source project of an LLM-based research assistant [53], the information within the reports is referenced to lower the chance of the model hallucinating. A hallucination occurs when an LLM generates a non-existent answer [58].

   The Process Researcher tool activates a separate Python virtual environment in which the GPT Researcher module runs, see Figure 14. A separate virtual environment is required because the GPT Researcher causes version conflicts with other components of the PoC. The GPT Researcher receives a query with five components: (1) the process component (step in the process), (2) the required analysis (e.g. bottleneck analysis), (3) the type of the process, (4) the sector, and (5) the organization. An example of a query used for this tool is: "What are common causes for process inefficiencies at the ‘Queued -> Completed’ step in the IT incidents managing process in the car manufacturing sector (at Volvo)?"

   The reports that are generated can have two levels of detail, standard (one GPT Researcher run, inspired by the Plan-and-Solve approach [59]) or detailed (multiple GPT Researcher runs, inspired by the STORM approach [60]). The approach that we use in this tool is an example of using RAG, as we let the agent use external information instead of relying on the pre-trained knowledge of the LLM.

4. **LLM:** A tool that allows the agent to ask simple questions to an LLM, e.g. to find the sector of an organization.

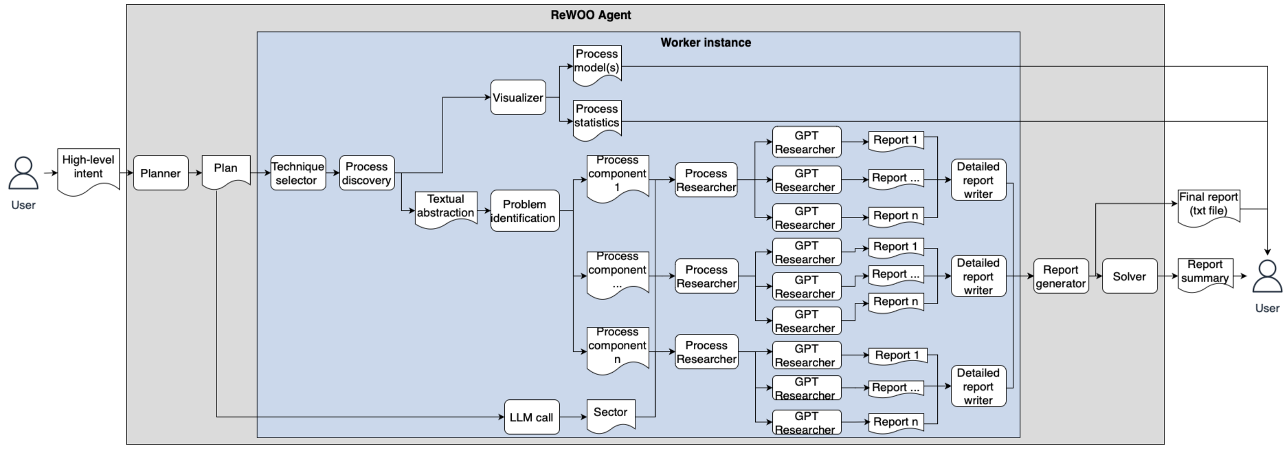


Figure 16: Architecture of the PoC during execution

5. **Human-in-the-loop:** This tool allows the agent to interact with the human user for input if the agent requires that, the response of the user is then used in the execution of the agent. With the inclusion of this tool, we allow for a human-in-the-loop option for the agent. For instance, the agent can ask questions to the user if the task is unclear. Making the agent more robust against unclear instructions from the user and lowering the risk of the system hallucinating.

6. **Report Generator:** This tool generates the final report, which aligns with the 'explanations' and 'improvements' steps of the framework (Figure 13) as this report includes potential explanations and improvements. The final report is returned as a separate text file because these files are too big to be returned to the user as the output of the agent.

Figure 16 shows an overview of the sequential steps of an execution of the agent using the mentioned tools, although this is not a static order of steps as the agent constructs a plan of action each time it is run. If the constructed plan is insufficient for solving the task, the agent can go back to the planner instance. The agent will adapt the plan and then execute this new plan. This type of looping within the agent is possible since the agent is based on a graph structure [61]. Figure 17 shows an overview of the graph structure of the agent, where it can be seen that the agent can go back to the 'plan' node after the tool execution has begun.

Then the system will spot the relevant process components from the generated textual abstraction of the process, based on the type of analysis that the user requested. Then for each identified relevant process component, the system will generate a (detailed) research report. In the generation of this report, the sector of the organization is also taken into account. As this can be valuable information for generating relevant domain knowledge about the process and its components. All explanations, improvements, and other forms of domain knowledge are written down into a research report about that specific process component. So for every identified process component, one research report is generated by the system. The system collects all the reports that it generates and adds these to one final report that is returned to the user (together with the requested visualizations).

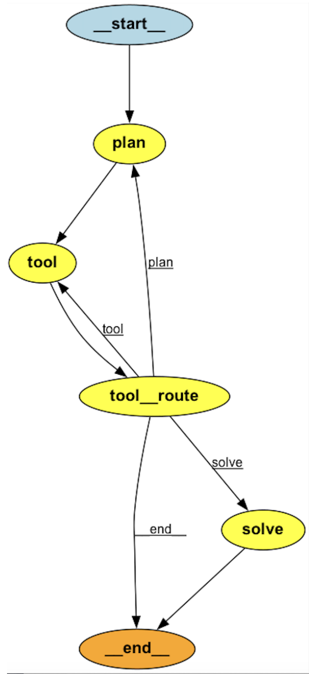


Figure 17: Visualized graph of an execution of the agent

6.2.2 LLMs

The goal of this PoC is to explore the potential of LLMs to supplement domain knowledge into process mining tools. For the PoC, we chose to use two different models of OpenAI, GPT-4 Turbo (gpt-4-1106-preview) and GPT-3.5 Turbo (gpt-3.5-turbo) [54]. We chose models of OpenAI for two main reasons. Firstly, the models from OpenAI are the most well-known and used LLMs out there currently. What this means is that these models are the most supported by tools and libraries for developing LLM-based systems. Secondly, the models from OpenAI are currently viewed as one of the best-performing LLMs in general [62]. Although, there are also other LLMs available that outperform the models from OpenAI on certain tests, for instance, Gemini Ultra beats GPT-4 on 'arithmetic reasoning' (GPT-4 finished second) [63].

GPT-4 Turbo is one of the most recently updated models from OpenAI, trained on knowledge until April 2023 [64]. Making it more capable than the previous generation model (GPT-3.5) and it has a larger token limit for input of 128,000 tokens. GPT-3.5 Turbo has a knowledge cutoff in September 2021 [65]. The token limit of this model is also lower than that of GPT-4 Turbo at 16,385 tokens [54]. The upside of GPT-3.5 Turbo compared to GPT-4 Turbo is the costs of using the model, currently for GPT-3.5 Turbo the costs of 1 million input tokens are \$0.50 and \$1.50 for 1 million output tokens [66]. For GPT-4 Turbo the costs of 1 million input tokens are \$10 and \$30 for 1 million output tokens.

Table 2 shows for each component in our PoC, which LLM we used. For simpler tasks that do not require a large amount of input or output tokens, we used GPT-3.5 Turbo. Because this model is cheaper to use than GPT-4 Turbo. For tasks that are more complex and that require a large amount of input or output tokens, we used GPT-4 Turbo. Within GPT Researcher both GPT-4 Turbo and GPT-3.5 Turbo are used, GPT-4 Turbo is used for more complex tasks and GPT-3.5 Turbo for simpler tasks.

Task (Figure 14)	Used LLM [54]
Agent foundation	OpenAI GPT-4 Turbo
LLM tool	OpenAI GPT-3.5 Turbo
Process analysis	OpenAI GPT-3.5 Turbo
Report tool	OpenAI GPT-4 Turbo
GPT Researcher	OpenAI GPT-4 Turbo & OpenAI GPT-3.5 Turbo

Table 2: Used LLM for each LLM-based part of the PoC

6.2.3 LangChain

LangChain is the technical basis for the PoC, LangChain is an open-source framework for developing LLM-based applications, with over five million monthly downloads [67]. We chose LangChain as the basis for this system as it is open-source, offers a large set of functions, simplifies the interaction with LLMs, and is regularly updated. Besides offering capabilities for building LLM-based applications, it also offers capabilities for monitoring and deploying these applications.

One of the modules within LangChain that we implemented in our PoC is LangGraph. LangGraph is a framework that offers the construction of multi-actor state LLM systems [61], using a graph structure. Using a graph structure for an LLM-based agent allows for the use of loops within the agent (Figure 17). In LangGraph states are an important concept, each graph execution a state is passed on between the nodes. Within each node, this state is updated and returned after execution [68]. Using a graph-based agent offers the mentioned multi-actor state capabilities and it gives the developer more control over the execution of the agent. Since the options that the agent has can be defined and connected via the edges of the graph, if required these edges can be made conditional. We chose to use LangGraph because a graph structure was required for implementing ReWOO and because it offers additional control over the execution of the agent.

LangSmith is an additional platform offered by LangChain to test and monitor LLM-based applications that are created with LangChain. It can give insights into the amount of LLM calls, token usage, success rate, and latency. Figure 18 shows a screenshot of the 'Monitor' tab of the LangSmith platform, where the graphs provide insights into the runs of the system. These insights are useful when trying to get a better understanding of your application and to for instance see how successful and efficient the executed runs are. Using LangSmith the claim of Xu et al. was observed for this system, showing a reduction in token usage and LLM calls when switching from ReAct to ReWOO [34]. Besides these statistics, it is also possible to get insights into the specific input and output of each LLM call and execution. This functionality of LangSmith can be helpful when debugging your application and seeing what instructions the model got at each step and what the respective output was.

6.2.4 GPT Researcher

The GPT Researcher module is used in the Process Researcher tool (Figure 14) of the agent for advanced searches. The GPT Researcher python library offers an autonomous agent that is designed to conduct advanced online searches, the code can be found on GitHub [53]. The final output of the

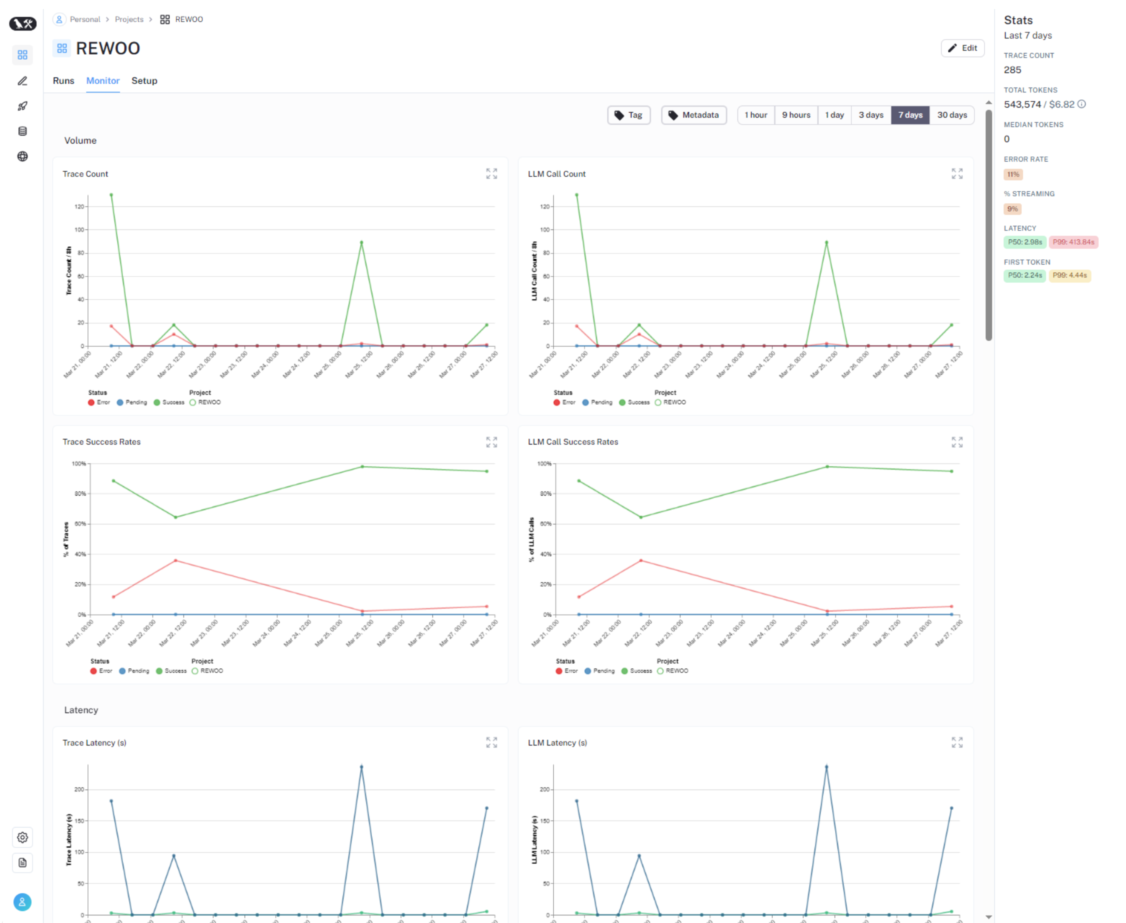
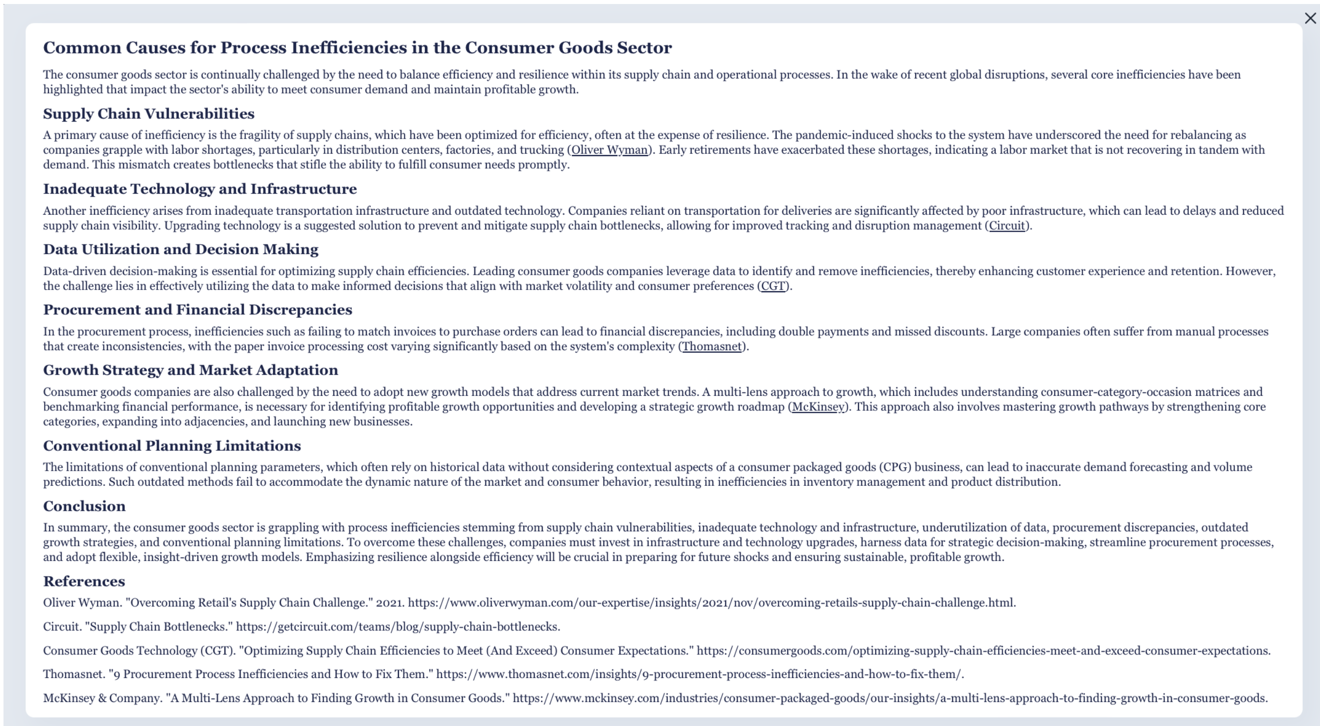


Figure 18: Screenshot of the LangSmith platform, from the 'Monitor' tab



Common Causes for Process Inefficiencies in the Consumer Goods Sector

The consumer goods sector is continually challenged by the need to balance efficiency and resilience within its supply chain and operational processes. In the wake of recent global disruptions, several core inefficiencies have been highlighted that impact the sector's ability to meet consumer demand and maintain profitable growth.

Supply Chain Vulnerabilities

A primary cause of inefficiency is the fragility of supply chains, which have been optimized for efficiency, often at the expense of resilience. The pandemic-induced shocks to the system have underscored the need for rebalancing as companies grapple with labor shortages, particularly in distribution centers, factories, and trucking (Oliver Wyman). Early retirements have exacerbated these shortages, indicating a labor market that is not recovering in tandem with demand. This mismatch creates bottlenecks that stifle the ability to fulfill consumer needs promptly.

Inadequate Technology and Infrastructure

Another inefficiency arises from inadequate transportation infrastructure and outdated technology. Companies reliant on transportation for deliveries are significantly affected by poor infrastructure, which can lead to delays and reduced supply chain visibility. Upgrading technology is a suggested solution to prevent and mitigate supply chain bottlenecks, allowing for improved tracking and disruption management (Circuit).

Data Utilization and Decision Making

Data-driven decision-making is essential for optimizing supply chain efficiencies. Leading consumer goods companies leverage data to identify and remove inefficiencies, thereby enhancing customer experience and retention. However, the challenge lies in effectively utilizing the data to make informed decisions that align with market volatility and consumer preferences (CGT).

Procurement and Financial Discrepancies

In the procurement process, inefficiencies such as failing to match invoices to purchase orders can lead to financial discrepancies, including double payments and missed discounts. Large companies often suffer from manual processes that create inconsistencies, with the paper invoice processing cost varying significantly based on the system's complexity (Thomasnet).

Growth Strategy and Market Adaptation

Consumer goods companies are also challenged by the need to adopt new growth models that address current market trends. A multi-lens approach to growth, which includes understanding consumer-category-occasion matrices and benchmarking financial performance, is necessary for identifying profitable growth opportunities and developing a strategic growth roadmap (McKinsey). This approach also involves mastering growth pathways by strengthening core categories, expanding into adjacencies, and launching new businesses.

Conventional Planning Limitations

The limitations of conventional planning parameters, which often rely on historical data without considering contextual aspects of a consumer packaged goods (CPG) business, can lead to inaccurate demand forecasting and volume predictions. Such outdated methods fail to accommodate the dynamic nature of the market and consumer behavior, resulting in inefficiencies in inventory management and product distribution.

Conclusion

In summary, the consumer goods sector is grappling with process inefficiencies stemming from supply chain vulnerabilities, inadequate technology and infrastructure, underutilization of data, procurement discrepancies, outdated growth strategies, and conventional planning limitations. To overcome these challenges, companies must invest in infrastructure and technology upgrades, harness data for strategic decision-making, streamline procurement processes, and adopt flexible, insight-driven growth models. Emphasizing resilience alongside efficiency will be crucial in preparing for future shocks and ensuring sustainable, profitable growth.

References

Oliver Wyman. "Overcoming Retail's Supply Chain Challenge." 2021. https://www.oliverwyman.com/our-expertise/insights/2021/nov/overcoming-retails-supply-chain-challenge.html.

Circuit. "Supply Chain Bottlenecks." https://getcircuit.com/teams/blog/supply-chain-bottlenecks.

Consumer Goods Technology (CGT). "Optimizing Supply Chain Efficiencies to Meet (And Exceed) Consumer Expectations." https://consumergoods.com/optimizing-supply-chain-efficiencies-meet-and-exceed-consumer-expectations.

Thomasnet. "9 Procurement Process Inefficiencies and How to Fix Them." https://www.thomasnet.com/insights/9-procurement-process-inefficiencies-and-how-to-fix-them/.

McKinsey & Company. "A Multi-Lens Approach to Finding Growth in Consumer Goods." https://www.mckinsey.com/industries/consumer-packaged-goods/our-insights/a-multi-lens-approach-to-finding-growth-in-consumer-goods.

Figure 19: Example of a 'standard' final report from GPT Researcher

GPT Researcher is a research report including references, Figure 19 shows an example for the query 'What are common causes for process inefficiencies in the consumer goods sector?'. We chose to use GPT Researcher because it provides references for the claims it makes, so they can be checked by the user. GPT Researcher goes one layer deeper than conventional search APIs for LLM-based agents and also analyzes the content of the found web page and not just their headline. This leads to the agent receiving more in-depth and detailed information and allows the agent to generate more in-depth domain knowledge based on the provided information.

The approach of the GPT Researcher was inspired by the Plan-and-Solve (PS) prompting [59]. GPT Researcher generates (sub) research questions for the query it receives and executes the sub-research questions in parallel, leading to a lower execution time for the agent. According to the makers, GPT Researcher is 85% faster than AutoGPT [69], a well-known open-source LLM-based autonomous agent [70]. To answer the (sub) research questions, the most relevant online sources are searched and information from these web pages is scraped and put into a report.

The mentioned information and example describe the 'standard' implementation of GPT Researcher, within the system, there is also an option to use a different one. This is the function that creates 'detailed reports', a more in-depth execution of the GPT Researcher inspired by the STORM paper of Shao et al. [60]. STORM stands for 'Synthesis of Topic Outlines through Retrieval and Multi-perspective Question Asking' [60]. These reports are around three times longer than the standard ones and contain around 20 references. Appendix B shows an example of a detailed report that uses the same query as the standard report in Figure 19. The downside of using these detailed reports is that this requires more execution time and tokens for completion. The detailed reports contain more information but can also contain information that is less relevant and specific to the question.

7 PoC results

In this section, we present the results of the evaluation of the PoC. We conducted two experiments to evaluate the output of the PoC. Firstly, an ablation experiment and secondly, a qualitative analysis. In the next two subsections, we present the results of both experiments.

7.1 Ablation experiment

We tested three different ablations to compare their outputs. The outputs that were generated are too large to fully display in this thesis, therefore, we present the first part of the outputs. The full outputs of our experiments can be found at https://github.com/maxvogt12/ReWOO_agent_for_PM in the 'Experiments' folder. The three ablations that we tested are: (1) the PM4PY LLM query functionality [7], (2) GPT Researcher [53], and (3) our PoC.

We asked all three systems what the audit risks are in an order-to-cash (O2C) process at Procter and Gamble (P&G), and what could be causing them. The dataset that we used can be found on our GitHub page, in the 'Event Logs' folder ('O2C.csv'). The data is about an O2C process but not specifically at P&G, we added this for testing purposes. The beginning of the output is shown in Figure 20 and our analysis of the output is shown in Table 3.

The PM4PY LLM query system was able to execute process discovery and problem identification but could not generate in-depth domain knowledge. For instance: "Manual order approval processes leading to delays in order processing.", this generated domain knowledge is superficial and not specific to this process, its component, or the organization.

We let the GPT Researcher generate a detailed report since that is also the default setting in our prototype, the beginning of this report is shown in Figure 20. GPT Researcher was able to generate more in-depth domain knowledge about O2C processes. However, because it cannot execute process discovery and problem identification, the domain knowledge is superficial in the sense that it is not focused on specific activities or parts of the O2C process. For instance: "Delayed payment processing is a critical risk in the cash application phase. For a company like Procter & Gamble, delays can disrupt the smooth operation of the order-to-cash cycle, affecting liquidity and the ability to reinvest in business operations. This can also lead to a poor customer experience, potentially damaging long-standing customer relationships and brand reputation". This output contains more in-depth information but the information is about the whole O2C process, GPT Researcher is not able to identify the components of this process and generate domain knowledge about those components.

Our PoC combines the strengths of these two systems, leading to in-depth domain knowledge that is focused on a specific part of the O2C process that was identified using process discovery and problem identification (Figure 13). The beginning of this report is shown on the right in Figure 20; we marked all information green that was specific for the process component that was passed along, compared to the 'generic' report of the GPT Researcher ablation (middle part of Figure 20).

In the output of our PoC, we see more detailed information, e.g. focusing on a specific activity in the process: "Segregation of duties is a key internal control that reduces the risk of errors and fraud. In the 'Send Purchase Order ->Receive Goods' step, the risk arises when the same individual is responsible for both ordering goods and receiving them. This lack of segregation can lead to unauthorized purchases or the acceptance of substandard goods. Procter & Gamble must ensure that different employees handle the procurement and receiving processes to mitigate this risk.

PM4PY LLM query:

Based on the DFG abstraction of the order-to-cash process at Procter and Gamble, some potential audit risks and their possible causes include:

1. Duplicate Orders:
- Cause: The frequency of "Order Validation -> Order Validation" is 13913, indicating a potential risk of duplicate orders being validated.
- Root Cause: Lack of system controls to prevent duplicate orders from entering the system.

2. Inaccurate Credit Checks:
- Cause: The frequency of "Customer Credit Check -> Customer Credit Check" is 27826, suggesting multiple credit checks being performed on the same customer.
- Root Cause: Inadequate integration between systems leading to repeated credit check requests.

3. Delayed Payments:
- Cause: The performance time for "Payment Received -> Goods Shipped" is 86400, indicating a potential delay in payments being processed.
- Root Cause: Inefficient payment processing systems or delays in reconciliation between payment received and goods shipped.

4. Manual Order Approvals:
- Cause: The performance time for "Order Validation -> Order Approval" is 320696.077, suggesting a lengthy approval process.
- Root Cause: Manual order approval processes leading to delays in order processing.

GPT Researcher without process context:

Audit Risks in the Order-to-Cash Process at Procter & Gamble

The order-to-cash (O2C) process is a critical business function within the Consumer Goods sector, encompassing all activities from order receipt through to cash collection. This report focuses on Procter & Gamble (P&G), a leader in the consumer goods industry, to identify and assess the inherent audit risks associated with their O2C process. Given the scale and complexity of operations at P&G, as detailed in their [financial and operational disclosures](https://pginvestor.com/financial-reporting/press-releases/news-details/2024/PG-Announces-Fiscal-Year-2024-Second-Quarter-Results/default.aspx), the O2C process is susceptible to various risks that could impact financial reporting and operational efficiency.

P&G's O2C cycle involves multiple critical steps, each presenting unique challenges and risks. These include order management, credit management, billing and invoicing, and cash application. The risks are compounded by factors such as global operation scales, diverse customer bases, and the integration of advanced technologies. Specific audit risks in P&G's O2C process include the accuracy and timeliness of invoice generation, the application of cash payments, and the management of credit extended to customers. Additionally, the company's reliance on digital platforms and automated systems introduces risks related to cybersecurity and data integrity.

This report will explore these risks in detail, considering the insights from industry experts and P&G's strategic responses to operational challenges as outlined in their [annual reports](https://us.pg.com/blogs/pg-earnings-second-quarter-2024/) and [EY's analysis of P&G's supply chain resilience](https://assets.ey.com/content/dam/ey-sites/ey-com/en_gl/topics/alliances/ey-pg-supply-chain-excellence-000421-22gbl.pdf). By examining these elements, we aim to provide a comprehensive audit risk assessment that will aid in enhancing the robustness of P&G's financial controls and operational strategies within the O2C process.

Table of Contents

- Audit Risks in Billing and Invoicing at Procter & Gamble
  - Risk of Inaccurate Billing Data
  - Compliance with Financial Reporting Standards
  - Fraud and Misappropriation Risks
  - System and Technology Risks
  - Impact of External Factors
- Cash Application and Payment Reconciliation Challenges at Procter & Gamble
  - Inefficiencies in Manual Cash Application Processes

Our PoC ("Send Purchase Order -> Receive Goods"):

Audit Risks in the Order to Cash Process: Procter & Gamble Case Study

The 'Send Purchase Order -> Receive Goods' step in the order to cash (O2C) process is a critical juncture in the supply chain of consumer goods companies like Procter & Gamble (P&G). This step involves the transmission of purchase orders to suppliers and the subsequent receipt of goods, which are foundational to maintaining inventory levels and ensuring product availability for end consumers. As a global leader in the consumer goods sector, P&G's operational efficiency and financial performance are heavily reliant on the effectiveness and reliability of this process. However, this step is not without its inherent risks, which can have significant financial, operational, and compliance implications.

In the context of P&G, a company with a vast and complex supply network ([P&G Supply Chain](https://us.pg.com/ethics-and-corporate-responsibility/supply-chain/)), the audit risks associated with the 'Send Purchase Order -> Receive Goods' step can be multifaceted. These risks include potential disruptions due to supply chain vulnerabilities, legal and financial liabilities stemming from purchase order inaccuracies, and the challenges of maintaining compliance with ethical sourcing and sustainability standards ([P&G Guidelines for Suppliers](https://pgsupplier.com/guidelines/pg-guidelines-for-suppliers)). Additionally, the company must navigate the risks of incorrect or delayed invoicing, which can lead to disputes or delayed payments, and the risk of errors in applying payments received from customers to their outstanding balances ([LinkedIn Article on O2C Risks](https://www.linkedin.com/pulse/order-cash-o2c-process-overview-key-risks-audit-muhammad-saad)).

Given the competitive nature of the consumer goods market, operational risks such as transportation failures and out-of-stock scenarios can have immediate repercussions on customer loyalty and market share ([Supply Chain Brain Article](https://www.supplychainbrain.com/blogs/1-think-tank/post/32970-retail-delivery-compliance-and-the-future-of-cpg-supply-chains)). Furthermore, the legal implications of purchase orders, such as potential breach of contract and misrepresentation, add a layer of complexity to the audit risk landscape ([Factwise Article on Purchase Orders](https://factwise.io/blog/post/legal-implications-purchase-orders)).

This report aims to provide a comprehensive analysis of the audit risks associated with the 'Send Purchase Order -> Receive Goods' step in P&G's O2C process. By examining the financial, operational, and compliance dimensions of these risks, we will offer insights into the potential impact on P&G's business and suggest strategies for risk mitigation. Our analysis is informed by the latest industry practices, regulatory requirements, and P&G's own supply chain management initiatives as of April 19, 2024.

Table of Contents
- Internal Control and Compliance Risks in the 'Send Purchase Order -> Receive Goods' Step
  - Segregation of Duties
  - Authorization of Purchase Orders
  - Verification of Goods Received
  - Accurate Recording of Transactions
  - Supplier Performance and Compliance
- Supply Chain and Delivery Risks in the 'Send Purchase Order -> Receive Goods' Step

Figure 20: Sample from the generated reports from the experiments

Table 3: Analysis of the output of the ablation experiment

Ablation	Analysis of output (see Figure 20 for output)
PM4PY LLM query	The textual abstraction (DFG) allows the LLM to analyze the process model and reason which activities in the process could form an audit risk based on the DFG (process domain knowledge). However, the explanations about the causes of these audit risks (ecosystem domain knowledge) are superficial compared to the other systems.
GPT Researcher	Generates more in-depth knowledge for the provided question. However, it is unaware of the components of the system and that makes the answers more superficial, the information is just about an O2C process but not about specific components within this process. GPT Researcher provides references for the information that it generates.
PoC	Executes process discovery and problem identification and can then provide relevant process components to the GPT Researcher. This allows for the generation of in-depth knowledge that is focused on the relevant parts of the process, in this case, the 'Send Purchase Order ->Receive Goods' step. Our PoC provides references for the information that it generates.

([Procter & Gamble Policies & Practices](https://us.pg.com/policies-and-practices/product-safety-and-compliance/))". In this example, we see information that is in-depth and focused on a specific process component and we see that documentation of the website of P&G was used as a reference.

The PoC can combine information about the organization with current trends to identify risks for P&G. The report mentioned that P&G is active in over 70 countries and therefore faces geopolitical risks like the war in the Middle East and the Houthi attacks on cargo ships. The used source is an example of RAG, the article is from February 2024 and the pre-trained knowledge of the GPT-4 model (currently) is cut off in December 2023 [71, 54]. The reports also include potential improvements for the observed audit risks in the process, e.g. that P&G can use the Third-Party Risk Management (TPRM) process and internal control frameworks that it has in place to mitigate some of the mentioned risks (including a reference to the website mentioning this framework).

For the ablation experiment, we ran another instance in the same way that we just presented. For this instance, we asked the same three systems (PM4PY, GPT Researcher, and our PoC) what the cyber security risks are in the IT incidents handling process at Volvo. We observed the same type of output from the three ablations as we described in Table 3. The PM4PY LLM query again was able to execute process mining but gave superficial information without references: "Unauthorized access: There is a risk of unauthorized access to sensitive information and systems when incidents are in the "Accepted" state, as well as when incidents move between different states without proper authorization.".

We observed that our PoC gave specific examples from the sector of the example organization, e.g.: "Connected vehicles and systems in the automotive industry are susceptible to various cyber threats, including malware, ransomware, and targeted attacks. For instance, infotainment systems, like Ford's Sync3, have been found to have vulnerabilities that could allow for remote code execution ([Telematicswire](https://www.telematicswire.net/upstream-releases-2024-automotive-cybersecurity-report/)).". Within these reports from the PoC, we see that the system presents potential risks first and then later discusses potential mitigation strategies for the risks. This aligns with the 'explanation' and 'improvement' components of our conceptual framework (Figure 13). An example of a mitigation strategy: "Utilizing guidelines from industry-specific organizations like Auto-ISAC provides tailored best practices for incident response, including roles and responsibilities, technical and non-technical processes, and self-evaluation after incident closure ([Auto-ISAC Best Practices](https://automotiveisac.com/best-practices/)).".

Furthermore, we see that the PoC mentions certain regulations that Volvo has to comply with, focusing specifically on cyber security: "Volvo adheres to automotive industry standards and regulatory requirements, such as UN155/156 set by UNECE WP.29, which mandates cybersecurity solutions throughout the vehicle lifecycle. Compliance with these standards is critical for maintaining market access and customer trust. ([UNECE WP.29](https://unece.org/))". The mentioned examples indicate the PoC's ability to generate domain knowledge. Another important factor of the PoC is that it is also able to execute process mining techniques and is able to zoom in on components of the process, which GPT Researcher is not able to do. One of the process components that the PoC focused on was the step from 'Queued' to 'Completed', e.g.: "Attackers often target known vulnerabilities in older software that are no longer supported by vendors. This can lead to incidents where the 'Queued' step involves a security breach, and the 'Completed' step requires not just resolving the immediate issue but also addressing the underlying security flaws, which can be complex and resource-intensive.". Without using process mining techniques, specifying the domain knowledge to this level would not have been possible.

Overall, the ablation experiment indicated that our PoC offers additional value over its two main components (separately). By being able to execute process mining and therefore finding the activities within the process. Selecting the relevant part of the process to analyze and then generating in-depth domain knowledge about the part of the process.

7.2 Qualitative analysis

As explained in Section 4.4.2, we evaluate how much of the information in the output of the PoC is specific information to the provided process component. We demonstrated this same approach (of measuring the amount of specific information in the report) in Figure 20, where all the green marked text on the right indicates specific instead of generic information. We marked all the information in the reports that referred to the provided component as specific information. The percentages in Table 4 and 5 represent the share of the report that contained this type of specific information.

In Table 4 we show the results of the qualitative analysis for all the reports that we generated and in Table 5 we show the results per generated report. In both tables, we ordered the rows with descending average scores, to be able to identify potential trends. We observe that the average scores for audit risk and environmental risk are the highest and the only ones that reach an average score of 90% or higher. The interesting thing about these two rows is that they are also the instances where the DFG technique was used as the textual abstraction. This might indicate that using the DFG technique leads to higher scores than the usage of the temporal profile or variants techniques. However, there could also be other reasons for the difference in the scores. For instance, there could be more online sources available about these problem identification types, the organization, or the process type.

If we look at the other four rows of Table 4 and 5, we see that there does not appear to be a trend for the usage of the variants or temporal profile technique and a potential effect on the average scores. If we specifically look at Table 5 and how these averages were calculated. We can observe that the three lowest-scoring rows are caused by one report scoring relatively low (under 66%). The other report for each of these three rows scores a lot higher, 79% or higher. This could indicate that these low scores for specific reports are outliers and that if more reports were generated the average scores (for these rows) would be higher. Although this is speculation and to verify this we would need to conduct a more in-depth evaluation of the PoC. Furthermore, we see that all averages are above 65%, so 65% of the report contained specific information. To see the full reports, visit https://github.com/maxvogt12/ReWOO_agent_for_PM, in the 'Experiments/Prototype_exp' folder.

While a large portion of the reports contains information specific to the process component, that does not mean that all of this information is useful for the process mining practitioner or other type of user. For instance, the system can sometimes still generate superficial information, although this information is specified to the process component and problem identification type. An example of this could be the following output: "Integrating the accounts payable system with procurement and financial systems ensures that all necessary data is readily available. This integration facilitates a smoother payment preparation process and helps maintain data consistency across systems.". This information is specified to the AP process and specifically the preparation of payment activity in this process. However, one could argue that this information is superficial and not specific enough for a user to gain new insights.

The percentages in Table 4 indicate the share of the information that was specified for the

Table 4: Compared results of our PoC

Problem identification	**Sector/ use case**	**Process type**	**Technique**	**Average**
Environmental risk	Retail	Purchase-to-Pay (P2P) [72]	DFG	92.58%
Audit risk	Consumer goods	Order-to-Cash (O2C) [73]	DFG	90.07%
IT & cyber risk	Automotive	IT incident handling [74]	Temporal profile	82.10%
Process inefficiencies	Manufacturing	Accounts Payable (AP) [75]	Variants	74.37%
Operational risk	Technology	Travel expenses [76]	Variants	68.64%
Regulatory risk	Financial services	Loan application [77]	Temporal profile	66.96%

Table 5: Results of our PoC per report

Problem identification	**Report 1**	**Report 2**	**Average**
Environmental risk	89.38%	95.77%	92.58%
Audit risk	90.84%	89.29%	90.07%
IT & cyber risk	81.51%	82.69%	82.10%
Process inefficiencies	82.85%	65.88%	74.37%
Operational risk	56.87%	80.40%	68.64%
Regulatory risk	79.49%	55.43%	66.96%

process component, which means that there is also a share that was not specified for such a component. Especially for the regulatory risk problem identification we observe that more than 30% of the report is information that is not specific. This could be caused by the type of problem identification, the sector, the process type, or the specific organization. For instance, if the system can find less information about an AP process compared to the O2C process, that can affect the quality of the output. The same holds for organizations, for large organizations with more online resources the system might be able to generate more specific output. Overall, the qualitative analysis indicated that within all reports across different problem identification types, over 65% of the content represented information (external ecosystem domain knowledge) specific to the process component(s). Based on these findings, we conclude that the PoC can generate domain knowledge about specific components (validating our first hypothesis). Therefore the system can at least partially take over the role of a human domain expert (second hypothesis). However, this does not mean that the PoC can fully replace a human domain expert.

8 Discussion

In this section, we first compare our hypotheses and research objectives with the findings we gathered. Secondly, we discuss the limitations of our work. Thirdly, based on these limitations we propose directions for future work.

8.1 Hypotheses & objectives

In Section 4, we presented our hypothesis and research objectives. The hypotheses are the following:

1. H_1*: An LLM-based system can provide domain knowledge about the context of a process to the user of a process mining tool.*

2. H_2*: An LLM-based system can partially take over the role of a human domain expert in the deployment of process mining.*

In Section 7, we presented the results from the conducted experiments. These results indicated that our PoC is capable of generating ecosystem domain knowledge about a process. Therefore we accept the first hypothesis. It should be noted that the generated domain knowledge is not necessarily complete and useful. For the second hypothesis, we established that our PoC can generate domain knowledge. Since one of the tasks of a human domain expert is to give domain knowledge, our system can partially take over this task. However, it would be interesting to study to which extent the PoC can take over the role of a human domain expert.

The presented research objectives in Section 4 are the following:

1. RO_1*: Develop a framework that describes how domain knowledge can be included in a process mining tool.*

2. RO_2*: Explore techniques that allow LLMs to produce useful output based on domain knowledge.*

3. RO_3*: Develop an LLM-based Proof-of-Concept (PoC) that can generate and supplement domain knowledge into a process mining tool.*

RO_1 states that we should develop a framework to include domain knowledge into a process mining tool. We presented our framework in Section 6.1.2. However, this framework is just one approach to do this and other approaches can also be used. For the developed PoC (RO_3) we used agent and prompt engineering techniques which we introduced in Section 2 (RO_2).

8.2 Limitations

Although our research presents a novel way to generate domain knowledge within process mining tools, it has its limitations. The limitations can be divided into two categories, limitations regarding the PoC and limitations regarding the evaluation of the PoC.

We will discuss the limitations regarding the PoC first. Our PoC works best for organizations or industries with a lot of online coverage, since the PoC searches for online sources about the specific process component, sector, and organization (external ecosystem domain knowledge). So if an organization has no online presence the results of the system might be less useful. If this type of system were to be used in a real-life scenario it could rely on documentation of the organization (internal ecosystem domain knowledge) to fill this gap in online presence. Often organizations have a lot of internal organizations that could be useful for the system to analyze, e.g. documentation about regulations and compliance. We demonstrated that the PoC can already rely on internal documentation, by using a document as the information source for the Technique Selector tool. Therefore, if the system is pointed toward the storage of internal documentation, it can analyze it and use it in the problem identification and for generating domain knowledge.

The PoC currently generates domain knowledge about a specific component of a process model, based on the names of the activities within this component. This works for event logs in which the activities have names that correctly describe the actual activity. However, there can also be event logs in which the activities are encoded or numbered instead of having descriptive names. In these cases, the PoC will not be able to directly generate specific domain knowledge about these activities and will generate domain knowledge about the entire process instead of a specific component. To fix this, one could implement a transformer in the system that translates the encoded activity names or numbers into descriptive names.

The PoC bases the generated domain knowledge on the outcomes of the process discovery and problem identification phases but does not allow the user to refine the problem identification phase, based on the generated domain knowledge. We did include this option in our conceptual vision (Figure 13) but we did not include it in the PoC.

For the evaluation of the PoC, we assessed three ablations of the system and conducted a qualitative analysis on multiple reports. These experiments did give insights into the added value of the prototype. However, they do not provide the complete picture, as the actual user experience is not taken into account in these two experiments. Furthermore, these experiments can be prone to bias, since they are based on our judgment, instead of including an independent benchmark to assess the quality of the generated output. This would allow us to get a better assessment of the quality of the output of the PoC. For the qualitative analysis, we were limited by a low number of publicly available process mining datasets. Therefore, we were not able to expand the results of the analyses shown in Table 4 and 5 by either looking at more problem identification types or generating more reports for the problem identification types that we already included.

The output of the PoC is rather large (often more than 10 pages) as we tried to generate as much relevant domain knowledge as possible to study the potential of LLMs for this task. However, for a business user, it might prove useful to investigate a filter for this large output and to just return the most useful output for a specific user. This can be realized by changing the prompts of the 'Solver' instance (Figure 16), to let it write a summary.

8.3 Future work

Now we will discuss directions for future work, based on our findings from this study. We see two main areas for future work, the development of an LLM-based system that generates domain knowledge and the evaluation of such a system.

8.3.1 System

The framework that we presented in our conceptual vision integrated the types of domain knowledge and process mining techniques (Figure 13). However, the PoC that we developed focused on specific components of this framework and did not include all its components. We can make a distinction between future work that focuses on expanding the domain knowledge capabilities and the other system capabilities.

For the domain knowledge, focusing more on the process knowledge could prove to be a useful direction. For instance, searching for data sources about process optimization and letting the agent analyze them. This could allow the PoC to become better at proposing process improvements. One could also look into other ways of integrating the domain knowledge into a process mining tool. For instance, using the approach proposed by Eichele et al. [50], injecting domain knowledge directly into the event logs. The quality of the generated domain knowledge depends on the quality of the inputs of the system, i.e. the information sources that it can find and analyze. Within these sources, we can make the distinction between internal and external sources. For internal sources, it would be interesting to study the potential of our PoC, if it would be given access to the internal documents and sources of an organization. And compare the results of the generated domain knowledge to the current output that we generated. For external, one might study how the system can find more and better sources for generating domain knowledge and look into the approaches that can be used to extract the knowledge from these sources. Finding the right information is important, however, this information also needs to be captured correctly and efficiently. Furthermore, a situation might arise where the generated domain knowledge and the event log or process model contradict each other. An example of such a situation could for instance be when a process is not compliant with a certain regulation. In this case, the process should be made compliant. Diving into these types of situations and mitigating the risks that come from them can make the system more robust.

Now we will discuss future work directions for the other system capabilities. To develop a system that completely aligns with our framework, the system should allow for looping between explanations or improvements and problem identification. This would give a more robust and interactive experience for the user. The user can zoom in on specific components in the process, get explanations and improvements about these components, and then refine the problem identification. For instance, switch to another component or go one layer deeper, and generate additional domain knowledge. Furthermore, the addition of additional functionalities for 'descriptive tasks' might prove useful. This would allow the user to ask questions to the system about the characteristics of the process model or the event log.

The world of LLMs is moving at a fast pace and new technologies are being developed constantly. Identifying new technologies that could improve this system and implementing one of them could prove to be beneficial for the value that our system offers. One new approach that is gaining popularity is DSPy. DSPy is an approach that allows the optimization of prompts for LLMs, removing the need for hardcoded prompts [78]. Currently, we apply LangChain to make the process of declaring and using prompts in the system simpler. However, using an approach like DSPy could allow us to make the development of our system even simpler and deliver a more robust final product. The downside of DSPy is that it requires training instances for the optimization of the prompts.

Another potential direction for future research is the evaluation of different LLMs, specifically for this type of system. As mentioned in Section 6.2.2, we currently use two different OpenAI models. With the current and future development of new LLMs, other LLMs might prove to yield a better performance for an LLM-based system for generating domain knowledge.

Besides studying the potential of using other LLMs, the usage of other process mining software might also prove to be beneficial. Currently, we use the PM4PY module in our PoC but the integration with a more advanced process mining suite (e.g. Celonis or Pegasystems) might result in better outcomes for our system. As these tools offer more in-depth process discovery and problem identification and if those analyses are improved, the quality of the generated domain knowledge might also improve.

8.3.2 Evaluation

To get a better understanding of the additional value of our proposed system the evaluation could be expanded. Specifically, looking into the experience users have when using such a system. For instance, letting a group of users use the system and observing them during their usage. Capturing their experience and collecting their thoughts afterward can allow us to get a better understanding of the expectations and experience of a user. This allows for a more complete validation of our findings and might give additional directions for future improvements of the system. By exploring the proposed directions for future work, we could be able to develop a system that offers even more value to the user and get a better understanding of how users experience and appreciate the system and the value that it offers.

9 Conclusion

This study explores the integration of LLMs into process mining tools, specifically focusing on the use of ReWOO-based agents to bridge the domain knowledge gap process mining practitioners face. We interviewed process mining practitioners to get insights into the current state of process mining and to see if and how LLMs could improve this. We distinguish three different types of domain knowledge within process mining and focus on the generation of ecosystem domain knowledge because this domain knowledge is currently not represented in process mining tools.

We present a conceptual vision of an LLM-based system that can generate domain knowledge and execute process mining techniques. Based on this conceptual vision, we developed a working PoC. Our PoC shows that LLMs can successfully generate and integrate domain-specific knowledge into process mining, potentially reducing the need for human experts and thus lowering barriers to effective process management. Through our experiments, we observe that our PoC provided more in-depth external ecosystem domain knowledge about specific process components compared to the other systems in the experiment.

Despite these promising outcomes, our study also highlights several challenges. Data availability and quality remain major hurdles, as the effectiveness of the generated insights heavily relies on the input data's comprehensiveness and accuracy. Looking ahead, we identify several directions for future work. Future studies can focus on further developing an LLM-based system for generating domain knowledge and the expansion of the evaluation of such a system. Incorporating a broader range of data sources could enrich the domain knowledge base, leading to more comprehensive support for process mining practitioners.

In conclusion, if we look at our research question *'How can LLM-based agents provide domain knowledge to help users of process mining tools understand and improve processes?'*. We conclude that LLM-based systems can generate domain knowledge by utilizing internal and external information sources (first hypothesis). Furthermore, we conclude that the generated domain knowledge is relevant for the users of process mining tools and that this knowledge can allow them to understand and improve processes. Therefore, an LLM-based system can (partially) take over the role of a human domain expert within process mining (second hypothesis).

References

[1] W. M. P. van der Aalst, “Process mining: discovering and improving spaghetti and lasagna processes,” in *2011 IEEE Symposium on Computational Intelligence and Data Mining (CIDM)*, 2011, pp. 1–7.

[2] F. Zerbato, P. Soffer, and B. Weber, “Process mining practices: evidence from interviews,” in *International Conference on Business Process Management*. Springer, 2022, pp. 268–285.

[3] T. Kampik, C. Warmuth, A. Rebmann, R. Agam, L. N. P. Egger, A. Gerber, J. Hoffart, J. Kolk, P. Herzig, G. Decker *et al.*, “Large process models: Business process management in the age of generative AI,” *CoRR*, vol. abs/2309.00900, 2023. [Online]. Available: http://dblp.uni-trier.de/db/journals/corr/corr2309.html#abs-2309-00900

[4] N. Martin, D. A. Fischer, G. D. Kerpedzhiev, K. Goel, S. J. Leemans, M. Röglinger, W. M. P. van der Aalst, M. Dumas, M. La Rosa, and M. T. Wynn, “Opportunities and challenges for process mining in organizations: results of a Delphi study,” *Business & Information Systems Engineering*, vol. 63, pp. 511–527, 2021.

[5] A. Mamudu, W. Bandara, M. T. Wynn, and S. J. Leemans, “A process mining success factors model,” in *International Conference on Business Process Management*. Springer, 2022, pp. 143–160.

[6] R. Andrews, M. T. Wynn, K. Vallmuur, A. H. M. ter Hofstede, and E. Bosley, “A comparative process mining analysis of road trauma patient pathways,” *International Journal of Environmental Research and Public Health*, vol. 17, no. 10, 2020. [Online]. Available: https://www.mdpi.com/1660-4601/17/10/3426

[7] A. Berti and M. S. Qafari, “Leveraging large language models (LLMs) for process mining (technical report),” *arXiv preprint arXiv:2307.12701*, 2023.

[8] A. Berti, S. J. van Zelst, and W. M. P. van der Aalst, “Process mining for python (PM4Py): Bridging the gap between process- and data science,” *CoRR*, vol. abs/1905.06169, 2019. [Online]. Available: http://arxiv.org/abs/1905.06169

[9] T. Smith. (2023) Process mining meets generative AI: Celonis rides industry wave to democratize core tech. https://accelerationeconomy.com/ai/process-mining-meets-generative-ai-celonis-rides-industry-wave-to-democratize-core-tech/. Accessed on 20-11-2023.

[10] Pegasystems. (2023) Pega launches Pega Process Mining with generative AI-ready APIs to enable continuous workflow optimization. https://www.pega.com/about/news/press-releases/pega-launches-pega-process-mining-generative-ai-ready-apis-enable. Accessed on 22-11-2023.

[11] K. Peffers, T. Tuunanen, M. A. Rothenberger, and S. Chatterjee, “A design science research methodology for information systems research,” *Journal of Management Information Systems*, vol. 24, no. 3, pp. 45–77, 2007. [Online]. Available: https://doi.org/10.2753/MIS0742-1222240302

[12] W. M. P. van der Aalst, "Process mining," *Communications of the ACM*, vol. 55, no. 8, pp. 76–83, 2012.

[13] M. Netjes, H. A. Reijers, and W. M. P. van der Aalst, "Supporting the BPM life-cycle with FileNet." in *EMMSAD*, 2006, pp. 135–146.

[14] W. M. P. van der Aalst, A. Adriansyah, A. K. A. De Medeiros, F. Arcieri, T. Baier, T. Blickle, J. C. Bose, P. van den Brand, R. Brandtjen, J. Buijs *et al.*, "Process mining manifesto," in *Business Process Management Workshops: BPM 2011 International Workshops, Clermont-Ferrand, France, August 29, 2011, Revised Selected Papers, Part I 9*. Springer, 2012, pp. 169–194.

[15] S. Goedertier, J. de Weerdt, D. Martens, J. Vanthienen, and B. Baesens, "Process discovery in event logs: An application in the telecom industry," *Applied Soft Computing*, vol. 11, no. 2, pp. 1697–1710, 2011, the Impact of Soft Computing for the Progress of Artificial Intelligence. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1568494610001055

[16] C. dos Santos Garcia, A. Meincheim, E. R. Faria Junior, M. R. Dallagassa, D. M. V. Sato, D. R. Carvalho, E. A. P. Santos, and E. E. Scalabrin, "Process mining techniques and applications – a systematic mapping study," *Expert Systems with Applications*, vol. 133, pp. 260–295, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0957417419303161

[17] S. Dunzer, M. Stierle, M. Matzner, and S. Baier, "Conformance checking: a state-of-the-art literature review," in *Proceedings of the 11th International Conference on Subject-Oriented Business Process Management*, ser. S-BPM ONE '19. New York, NY, USA: Association for Computing Machinery, 2019. [Online]. Available: https://doi.org/10.1145/3329007.3329014

[18] S. Smirnov, H. A. Reijers, M. Weske, and T. Nugteren, "Business process model abstraction: a definition, catalog, and survey," *Distributed and Parallel Databases*, vol. 30, pp. 63–99, 2012.

[19] A. Kalenkova, W. M. P. van der Aalst, I. Lomazova, and V. Rubin, "Process mining using BPMN: Relating event logs and process models // process mining using BPMN. relating event logs and process models," *Software and Systems Modeling*, pp. 1–30, 01 2015.

[20] A. Berti, S. van Zelst, and D. Schuster, "PM4Py: A process mining library for python," *Software Impacts*, vol. 17, p. 100556, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2665963823000933

[21] E. Kasneci, K. Seßler, S. Küchemann, M. Bannert, D. Dementieva, F. Fischer, U. Gasser, G. Groh, S. Günnemann, E. Hüllermeier *et al.*, "ChatGPT for good? on opportunities and challenges of large language models for education," *Learning and individual differences*, vol. 103, p. 102274, 2023.

[22] Y. Wang, S. Si, D. Li, M. Lukasik, F. Yu, C.-J. Hsieh, I. S. Dhillon, and S. Kumar, "Preserving in-context learning ability in large language model fine-tuning," *arXiv preprint arXiv:2211.00635*, 2022.

[23] K. Busch, A. Rochlitzer, D. Sola, and H. Leopold, "Just tell me: Prompt engineering in business process management," in *International Conference on Business Process Modeling, Development and Support*. Springer, 2023, pp. 3–11.

[24] J. White, Q. Fu, S. Hays, M. Sandborn, C. Olea, H. Gilbert, A. Elnashar, J. Spencer-Smith, and D. C. Schmidt, "A prompt pattern catalog to enhance prompt engineering with ChatGPT," *arXiv preprint arXiv:2302.11382*, 2023.

[25] J. Zamfirescu-Pereira, R. Y. Wong, B. Hartmann, and Q. Yang, "Why Johnny can't prompt: How non-AI experts try (and fail) to design LLM prompts," in *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, ser. CHI '23. New York, NY, USA: Association for Computing Machinery, 2023. [Online]. Available: https://doi.org/10.1145/3544548.3581388

[26] Q. Dong, L. Li, D. Dai, C. Zheng, Z. Wu, B. Chang, X. Sun, J. Xu, and Z. Sui, "A survey for in-context learning," *arXiv preprint arXiv:2301.00234*, 2022.

[27] J. Wei, X. Wang, D. Schuurmans, M. Bosma, F. Xia, E. Chi, Q. V. Le, D. Zhou *et al.*, "Chain-of-thought prompting elicits reasoning in large language models," *Advances in Neural Information Processing Systems*, vol. 35, pp. 24 824–24 837, 2022.

[28] P. Lewis, E. Perez, A. Piktus, F. Petroni, V. Karpukhin, N. Goyal, H. Küttler, M. Lewis, W. tau Yih, T. Rocktäschel, S. Riedel, and D. Kiela, "Retrieval-augmented generation for knowledge-intensive NLP tasks," *Advances in Neural Information Processing Systems*, vol. 33, pp. 9459–9474, 2020.

[29] X. Chen, L. Li, N. Zhang, X. Liang, S. Deng, C. Tan, F. Huang, L. Si, and H. Chen, "Decoupling knowledge from memorization: Retrieval-augmented prompt learning," *Advances in Neural Information Processing Systems*, vol. 35, pp. 23 908–23 922, 2022.

[30] L. Monigatti. (2023) Retrieval-Augmented Generation (RAG): From theory to LangChain implementation. https://towardsdatascience.com/retrieval-augmented-generation-rag-from-theory-to-langchain-implementation-4e9bd5f6a4f2. Accessed on 28-05-2024.

[31] Z. Xi, W. Chen, X. Guo, W. He, Y. Ding, B. Hong, M. Zhang, J. Wang, S. Jin, E. Zhou *et al.*, "The rise and potential of large language model based agents: A survey," *arXiv preprint arXiv:2309.07864*, 2023.

[32] L. Wang, C. Ma, X. Feng, Z. Zhang, H. Yang, J. Zhang, Z. Chen, J. Tang, X. Chen, Y. Lin *et al.*, "A survey on large language model based autonomous agents," *Frontiers of Computer Science*, vol. 18, no. 6, p. 186345, 2024.

[33] S. Yao, J. Zhao, D. Yu, N. Du, I. Shafran, K. R. Narasimhan, and Y. Cao, "ReAct: Synergizing reasoning and acting in language models," in *The Eleventh International Conference on Learning Representations*, 2023. [Online]. Available: https://openreview.net/forum?id=WE_vluYUL-X

[34] B. Xu, Z. Peng, B. Lei, S. Mukherjee, Y. Liu, and D. Xu, "ReWOO: Decoupling reasoning from observations for efficient augmented language models," *arXiv preprint arXiv:2305.18323*, 2023.

[35] V. Venkatesh, J. Y. Thong, and X. Xu, "Unified theory of acceptance and use of technology: A synthesis and the road ahead," *Journal of the association for Information Systems*, vol. 17, no. 5, pp. 328–376, 2016.

[36] V. Venkatesh, M. G. Morris, G. B. Davis, and F. D. Davis, "User acceptance of information technology: Toward a unified view," *MIS quarterly*, pp. 425–478, 2003.

[37] Fortune Business Insights. (2023) Process mining software market. https://www.fortunebusinessinsights.com/process-mining-software-market-104792. Accessed on 20-04-2024.

[38] R. Syed, S. J. Leemans, R. Eden, and J. A. Buijs, "Process mining adoption: a technology continuity versus discontinuity perspective," in *Business Process Management Forum: BPM Forum 2020, Seville, Spain, September 13–18, 2020, Proceedings 18*. Springer, 2020, pp. 229–245.

[39] M. L. Tushman and P. Anderson, "Technological discontinuities and organizational environments," in *Organizational Innovation*. Routledge, 2018, pp. 345–372.

[40] L. Reinkemeyer, "Status and future of process mining: from process discovery to process execution," in *Process Mining Handbook*. Springer, 2022, pp. 405–415.

[41] N. Klievtsova, J. V. Benzin, T. Kampik, J. Mangler, and S. Rinderle-Ma, "Conversational process modelling: state of the art, applications, and implications in practice," in *International Conference on Business Process Management*. Springer, 2023, pp. 319–336.

[42] K. Busch, A. Rochlitzer, D. Sola, and H. Leopold, "Just tell me: Prompt engineering in business process management," in *International Conference on Business Process Modeling, Development and Support*. Springer, 2023, pp. 3–11.

[43] A. Berti, D. Schuster, and W. M. P. van der Aalst, "Abstractions, scenarios, and prompt definitions for process mining with LLMs: A case study," *ArXiv*, vol. abs/2307.02194, 2023. [Online]. Available: https://api.semanticscholar.org/CorpusID:259342720

[44] T. Kojima, S. S. Gu, M. Reid, Y. Matsuo, and Y. Iwasawa, "Large language models are zero-shot reasoners," *Advances in neural information processing systems*, vol. 35, pp. 22 199–22 213, 2022.

[45] U. Jessen, M. Sroka, and D. Fahland, "Chit-chat or deep talk: Prompt engineering for process mining." arXiv.org, WorkingPaper, 2023.

[46] H. Kourani, A. Berti, D. Schuster, and W. M. P. van der Aalst, "ProMoAI: Process modeling with generative AI," *arXiv preprint arXiv:2403.04327*, 2024.

[47] M. Grohs, L. Abb, N. Elsayed, and J. R. Rehse, "Large language models can accomplish business process management tasks," *ArXiv*, vol. abs/2307.09923, 2023. [Online]. Available: https://api.semanticscholar.org/CorpusID:259983054

[48] A. Rebmann, F. D. Schmidt, G. Glavaš, and H. van der Aa, "Evaluating the ability of LLMs to solve semantics-aware process mining tasks," *arXiv preprint arXiv:2407.02310*, 2024.

[49] SAP. (2023) Process AI: Generative AI for the process world. https://www.signavio.com/process-ai/. Accessed on 05-03-2024.

[50] S. Eichele, K. Hinkelmann, and M. Spahic-Bogdanovic, "Ontology-driven enhancement of process mining with domain knowledge," in *AAAI 2023 Spring Symposium on Challenges Requiring the Combination of Machine Learning and Knowledge Engineering*, 2023.

[51] C. von Soest, "Why do we speak to experts? reviving the strength of the expert interview method," *Perspectives on Politics*, vol. 21, no. 1, pp. 277–287, 2023.

[52] S. Sheikholeslami, M. Meister, T. Wang, A. H. Payberah, V. Vlassov, and J. Dowling, "AutoAblation: Automated parallel ablation studies for deep learning," in *Proceedings of the 1st Workshop on Machine Learning and Systems*, ser. EuroMLSys '21. New York, NY, USA: Association for Computing Machinery, 2021, p. 55–61. [Online]. Available: https://doi.org/10.1145/3437984.3458834

[53] A. Elovic. (2024) GPT Researcher. https://github.com/assafelovic/gpt-researcher?ref=blog.langchain.dev. Accessed on 26-04-2024.

[54] OpenAI. (2024) Models. https://platform.openai.com/docs/models/overview. Accessed on 28-04-2024.

[55] Celonis. (2023) PQL - Process Query Language. https://docs.celonis.com/en/pql---process-query-language.html. Accessed on 12-03-2024.

[56] Q. Zhang, M. Irfan, M. A. O. Khattak, X. Zhu, and M. Hassan, "Lean Six Sigma: a literature review," *Interdisciplinary Journal of Contemporary research in business*, vol. 3, no. 10, pp. 599–605, 2012.

[57] M. Vogt. (2024) ReWOO_agent_for_PM. https://github.com/maxvogt12/ReWOO_agent_for_PM. Accessed on 08-07-2024.

[58] J.-Y. Yao, K.-P. Ning, Z.-H. Liu, M.-N. Ning, and L. Yuan, "LLM lies: Hallucinations are not bugs, but features as adversarial examples," *arXiv preprint arXiv:2310.01469*, 2023.

[59] L. Wang, W. Xu, Y. Lan, Z. Hu, Y. Lan, R. K.-W. Lee, and E.-P. Lim, "Plan-and-Solve prompting: Improving zero-shot chain-of-thought reasoning by large language models," in *Annual Meeting of the Association for Computational Linguistics*, 2023. [Online]. Available: https://api.semanticscholar.org/CorpusID:258558102

[60] Y. Shao, Y. Jiang, T. A. Kanell, P. Xu, O. Khattab, and M. S. Lam, "Assisting in writing Wikipedia-like articles from scratch with large language models," *arXiv preprint arXiv:2402.14207*, 2024.

[61] M. Zhuge, W. Wang, L. Kirsch, F. Faccio, D. Khizbullin, and J. Schmidhuber, "Language agents as optimizable graphs," *arXiv preprint arXiv:2402.16823*, 2024.

[62] Geeks for Geeks. (2024) Top 20 LLM models. https://www.geeksforgeeks.org/top-20-llm-models/. Accessed on 03-05-2024.

[63] S. Minaee, T. Mikolov, N. Nikzad, M. Chenaghlu, R. Socher, X. Amatriain, and J. Gao, "Large language models: A survey," *arXiv preprint arXiv:2402.06196*, 2024.

[64] OpenAI. (2024) GPT-4 turbo in the OpenAI API. https://help.openai.com/en/articles/8555510-gpt-4-turbo-in-the-openai-api. Accessed on 27-04-2024.

[65] ——. (2024) GPT-3.5 turbo updates. https://help.openai.com/en/articles/8555514-gpt-3-5-turbo-updates. Accessed on 27-04-2024.

[66] ——. (2024) Pricing. https://openai.com/pricing. Accessed on 27-04-2024.

[67] LangChain Inc. (2024) Applications that can reason. powered by LangChain. https://www.langchain.com. Accessed on 02-04-2024.

[68] LangChain. (2024) LangGraph. https://python.langchain.com/docs/langgraph. Accessed on 20-04-2024.

[69] A. Elovic. (2023) How we built GPT Researcher. https://docs.tavily.com/blog/building-gpt-researcher. Accessed on 19-04-2024.

[70] AutoGPT. (2023) Auto-GPT. https://news.agpt.co. Accessed on 20-04-2024.

[71] D. Kuketz and H. Schwarz. (2024) Top 10 operational risks for 2024. https://sphera.com/spark/top-10-operational-risks-for-2024/. Accessed on 02-05-2024.

[72] C. Alvite. (2023) purchase-to-pay event log. https://drive.google.com/file/d/1COTWOz1LD76imYRPbtRw_M6EJwPV3HTV/view?usp=share_link. Accessed on 08-05-2024.

[73] ——. (2023) order-to-cash event log. https://drive.google.com/file/d/1PWJNvw3fkUIExrr01eqkhSFRjkZkWVs9/view?usp=share_link. Accessed on 08-05-2024.

[74] W. Steeman. (2013) BPI challenge 2013, incidents. https://data.4tu.nl/articles/dataset/BPI_Challenge_2013_incidents/12693914/1. Accessed on 05-04-2024.

[75] C. Alvite. (2023) Ap event log. https://drive.google.com/file/d/1kGfZt60iRkgSORZWwfYh4qiggD1HoCxm/view?usp=share_link. Accessed on 08-05-2024.

[76] B. van Dongen. (2020) BPI challenge 2020: Prepaid travel costs. https://data.4tu.nl/articles/dataset/BPI_Challenge_2020_Prepaid_Travel_Costs/12696722/1. Accessed on 05-04-2024.

[77] ——. (2017) BPI challenge 2017. https://data.4tu.nl/articles/dataset/BPI_Challenge_2017/12696884. Accessed on 05-04-2024.

[78] O. Khattab, A. Singhvi, P. Maheshwari, Z. Zhang, K. Santhanam, S. Vardhamanan, S. Haq, A. Sharma, T. T. Joshi, H. Moazam *et al.*, "DSPy: Compiling declarative language model calls into self-improving pipelines," *arXiv preprint arXiv:2310.03714*, 2023.

A Interview questions

Interview questions list

Background:

1. What is your academic background in terms of formal education?
2. What is your current job function?
3. How long is your overall professional experience?

Process mining & domain knowledge:

4. What and how long is your overall experience with process mining?
5. What are (in your opinion) barriers in current process mining that limit the adoption of process mining? In terms of broadening to a business audience, in terms of more repeat use and/or in terms of increased coverage in types of processes and domains.
   a. Why?
6. Do you think that domain knowledge about the context of the process and its organization is important?
   a. Why?
7. What type of domain knowledge do you think is important? Knowledge about process mining? About process optimization in general? Or about the business domain the process is about?
8. Do you see this knowledge well represented in process mining tools?
   a. Can you give an example?
9. Can having too much domain knowledge about the process be a potential disadvantage?
   a. Why?

Generative AI:

10. What and how long is your overall experience with generative AI?
11. How often do you use generative AI?
    a. When and which system?
12. How good or bad has your experience and interactions with generative AI been?
    a. Why?
13. Are there specific use cases or challenges where you believe generative AI could add significant value in process mining?
14. How could the current generative AI systems (that you have used) improve?
    a. Why?

Prototype/ technology acceptance:

15. Do you always accept the output provided by generative AI systems?
    a. Why?
16. In your opinion, what factors influence the acceptance of new technologies, especially in the context of process mining or generative AI?

Closing:

17. Do you have any additional recommendations regarding the development of the proposed prototype?

18. Are there specific challenges or considerations you think should be prioritized during the prototype development process?
19. Is there anything else you would like to share regarding process mining, generative AI, or technology acceptance?
20. Are you available for follow-up questions?

B Detailed research report

Process Inefficiencies in the Consumer Goods Sector

The consumer goods sector, a dynamic and integral component of the global economy, is currently navigating a complex landscape marked by rapid technological advancements, shifting market dynamics, and evolving consumer preferences. Despite the sector's agility and resilience, process inefficiencies persist, often undermining the potential for growth and profitability. This report aims to dissect the common causes of such inefficiencies, drawing insights from industry leaders and authoritative sources.

In 2023, the consumer goods industry faced significant headwinds, including [dramatic cost inflation](https://www.bain.com/insights/consumer-products-report-2024-resetting-the-growth-agenda/), tight labor markets, high capacity utilization, and the end of ultra-low interest rates. These factors collectively constrained operational flexibility and highlighted the diminishing returns of traditional cost-saving measures. As overhead costs have been driven down since 2008, the room for further reductions has narrowed, necessitating a strategic pivot.

The [Bain & Company's Consumer Products Report](https://www.bain.com/insights/consumer-products-report-2024-resetting-the-growth-agenda/) underscores the importance of formulating new strategies to address these challenges. Similarly, the [Consumer Goods Industry Trends 2024 Report](https://www.dynamicyield.com/wp-content/uploads/2024/03/Consumer_goods_industry_trends_2024_Report_V2.pdf) by Dynamic Yield emphasizes the evolving consumer shopping habits and the need for consumer goods companies to adapt to ongoing shifts in purchasing behaviors and spending willingness.

Moreover, the industry is grappling with the imperative of digital transformation. As per the [Harvard Business Review](https://www.dynamicyield.com/wp-content/uploads/2024/03/Consumer_goods_industry_trends_2024_Report_V2.pdf), the consumer goods sector lags in tech capability and architecture, necessitating substantial investment in talent and analytics to remain competitive. The rise of [Generative AI](https://www.dynamicyield.com/wp-content/uploads/2024/03/Consumer_goods_industry_trends_2024_Report_V2.pdf) and other digital tools offers opportunities for innovation but also presents challenges in terms of integration and execution.

Supply chain management remains a critical area where inefficiencies can have a cascading effect on the entire operation. [McKinsey & Company](https://www.mckinsey.com/industries/consumer-packaged-goods/our-insights/supply-chain-4-0-in-consumer-goods) highlights the potential of Supply Chain 4.0 to dramatically reduce operational costs and lost sales while decreasing inventories. However, achieving these benefits requires a strategic approach to data analytics, demand planning, and network optimization.

In conclusion, the consumer goods sector is at a crossroads, where addressing process inefficiencies is not just about cost-cutting but about reimagining business models and embracing digital and sustainable practices. This report will delve into the multifaceted causes of inefficiencies, exploring solutions that can drive the sector towards a more resilient and prosperous future.

Table of Contents

Challenges and Trends in the Consumer Goods Sector

Supply Chain Challenges

The consumer goods sector continues to face significant supply chain challenges, which are a common cause of process inefficiencies. Recent attacks on oceanic freight lines in the Red Sea have led companies to reevaluate their supplier bases, with many transitioning away from Asian manufacturing hubs to alternatives like Vietnam and India ([RSM US](https://rsmus.com/insights/industries/consumer-goods/consumer-goods-outlook.html)). However, these transitions are complex and costly, and geopolitical risks necessitate ongoing reviews of manufacturing strategies. Elevated operating costs also challenge suppliers and manufacturers, with geopolitics, labor tensions, and freight demand fluctuations posing additional risks ([Supply Chain Dive](https://www.supplychaindive.com/news/supply-chain-trends-outlook-2024/706096/)).

Technological Integration and Data Analytics

Investment in predictive analytical tools and data analytics platforms is critical for addressing inefficiencies in the consumer goods sector. Brands are dealing with an average of 976 siloed systems, which hampers the efficiency of operations. By preparing data for generative AI, consumer goods brands can transform marketing, commerce, service, and operations, leading to improved process efficiencies ([Salesforce](https://www.salesforce.com/blog/consumer-goods-industry-trends/)).

Market Dynamics and Consumer Behavior

Value-driven consumption is reshaping the marketplace, with consumers increasingly making choices based on a brand's values and ethical practices. This shift in consumer behavior emphasizes the importance of sustainability, social responsibility, and ethical practices in purchasing decisions, which consumer goods companies must adapt to ([Kadence International](https://www.kadence.com/en-us/15-consumer-trends-to-watch-in-2024/)). Additionally, the rise of social media influence on consumer behavior is transforming how brands interact with consumers, particularly in sectors like Advertising, Media, and markets targeting Kids and Youth.

E-commerce Innovation

E-commerce innovation continues to shape the landscape of B2B transactions, redefine consumer experiences in the CPG sector, and challenge traditional financial services models. The businesses that embrace these innovations and adapt to the evolving digital marketplace will be well-positioned to thrive. This trend towards home-centric products and services is driving innovation across multiple industries, and companies that can effectively cater to consumers' home-based needs and preferences will likely see sustained growth and success ([Kadence International](https://www.kadence.com/en-us/15-consumer-trends-to-watch-in-2024/)).

Mergers and Acquisitions (M&A) Dynamics

On the M&A front, 2023 saw a decline in deal activity due to economic uncertainty and valuation gaps. While increased earnout arrangements and add-on acquisitions continue into 2024, investor caution persists. Sectors catering to cost-conscious consumers, like mass-market consumer goods, face ongoing stress, while home services and beauty brands remain attractive to investors seeking strong multiples ([RSM US](https://rsmus.com/insights/industries/consumer-goods/consumer-goods-outlook.html)).

Digital Transformation and Investment in Consumer Goods

Digital Transformation: A Catalyst for Efficiency

Digital transformation in the consumer goods sector is a strategic integration of digital technologies into all aspects of business, from manufacturing and supply chains to sales and customer engagement. By adopting technologies such as AI, IoT, and blockchain, companies aim to enhance business processes, including inventory management, manufacturing efficiency, and product customization. For instance, the use of Opcenter RD&L integrates laboratory information management systems (LIMS) to drive innovation

efficiency
([KPMG](https://kpmg.com/xx/en/home/insights/2024/03/kpmg-global-tech-report-2023-consumer-and-retail-sector-insights.html)).

Overcoming Inefficiencies with Digital Solutions

Consumer goods companies are turning to digital tools to address inefficiencies in manufacturing processes. Digital tools applied to lean transformations and advanced analytics optimize specific manufacturing processes, leading to improved efficiency and productivity ([McKinsey & Company](https://www.mckinsey.com/industries/consumer-packaged-goods/our-insights/digital-innovation-in-consumer-goods-manufacturing)).

The Role of Data in Streamlining Operations

Investment in data strategies and tools can lead to increased visibility, faster product launches, and cost reduction. Data-driven insights are crucial for business operations and growth, affecting systems, processes, workflows, and culture at all levels. Emphasizing agility, innovation, and resilience, consumer goods companies can respond more swiftly to market volatility, making digital transformation non-negotiable
([Layerise](https://www.layerise.com/resources/blog/post/strategies-to-accelerate-digital-transformation-in-consumer-goods-companies)).

Enhancing Customer Engagement through Technology

Digital transformation investments have significantly improved customer engagement for consumer and retail companies. For example, 23 percent of companies report that their investments have performed above expectations in this area. This improvement in engagement can lead to reduced friction in customer experiences and increased satisfaction ([KPMG](https://kpmg.com/xx/en/home/insights/2024/03/kpmg-global-tech-report-2023-consumer-and-retail-sector-insights.html)).

Addressing Legacy Systems and Tech Debt

Consumer and retail executives are less confident in their ability to enhance employee satisfaction and wellbeing using their current tech stack. This could be due to complex ecosystems and legacy tech debt that retail staff often have to navigate. Addressing these issues is essential for improving employee productivity and reducing inefficiencies caused by high employee attrition rates
([KPMG](https://kpmg.com/xx/en/home/insights/2024/03/kpmg-global-tech-report-2023-consumer-and-retail-sector-insights.html)).

In summary, digital transformation and investment in consumer goods are pivotal for addressing process inefficiencies. By leveraging digital tools, data strategies, and customer engagement technologies, companies can streamline operations, enhance productivity, and improve overall efficiency. Addressing legacy systems and tech debt is also crucial for maximizing the benefits of digital transformation.

Efficiency in Supply Chain Management

Supply Chain Visibility and Tracking
In the consumer goods sector, a lack of supply chain visibility is a common cause of inefficiency. Without clear visibility, companies struggle to track product flow, leading to inventory mismanagement and delayed response to supply chain disruptions. Advanced tracking technologies and integration of IoT devices provide real-time data, enabling better decision-making and more efficient supply chain operations ([NetSuite](https://www.netsuite.com/portal/resource/articles/inventory-management/supply-chain-efficiency.shtml)).

Inventory Management and Demand Forecasting
Inefficient inventory management, often due to poor demand forecasting, results in either excess stock or stockouts. Companies are turning to AI and machine learning to enhance demand forecasting accuracy, thereby optimizing inventory levels and reducing holding costs. This shift towards intelligent systems is redefining inventory management practices in the consumer goods sector ([Forbes](https://www.forbes.com/sites/paulnoble/2024/03/11/the-future-of-supply-chain-technology-a-shift-toward-intelligent-systems/)).

Supply Chain Process Integration
Process inefficiencies often stem from a lack of integration across the supply chain. Siloed operations prevent the seamless flow of information and goods, leading to delays and increased costs. By adopting integrated supply chain management systems, companies can streamline processes, improve collaboration, and enhance overall efficiency ([KPMG](https://kpmg.com/xx/en/home/insights/2023/12/supply-chain-trends-2024.html)).

Sustainable and Agile Supply Chain Practices
Sustainability initiatives, while beneficial in the long term, can introduce short-term inefficiencies if not properly managed. However, companies that invest in sustainable practices, such as fleet electrification and green warehousing, often find that these initiatives lead to greater efficiency through reduced waste and improved resource management. Additionally, building an agile supply chain that can quickly adapt to changes is crucial for maintaining efficiency in a volatile market ([ASCM](https://www.ascm.org/globalassets/ascm_website_assets/docs/top-10-trends-report-2024.pdf)).

Talent Development and Workforce Management
A skilled workforce is essential for efficient supply chain management. The consumer goods sector faces challenges in attracting and retaining talent with the necessary skills to manage modern, technologically advanced supply chains. Companies are addressing this gap through targeted hiring, staff training, and the use of collaborative robots to support human workers, thereby improving process efficiency ([NetSuite](https://www.netsuite.com/portal/resource/articles/erp/supply-chain-best-practices.shtml)).

References

- https://www.mckinsey.com/industries/consumer-packaged-goods/our-insights/the-consumer-sector-in-2030-trends-and-questions-to-consider
- https://www.mckinsey.com/capabilities/strategy-and-corporate-finance/our-insights/the-next-frontier-in-consumer-goods-digitally-enabled-innovation
- https://www2.deloitte.com/us/en/pages/consumer-business/articles/consumer-products-industry-outlook.html
- https://www.spglobal.com/_assets/documents/ratings/research/101594285.pdf
- https://www.econlib.org/library/Topics/College/marketfailures.html
- https://www.apqc.org/resource-library/resource-collection/2024-supply-chain-priorities-and-challenges
- https://www.supplychaindive.com/news/supply-chain-trends-outlook-2024/706096/
- https://supplychaindigital.com/technology/gartners-supply-chain-tech-trends-2024
- https://www.mckinsey.com/industries/retail/our-insights/state-of-grocery-europe-2024-signs-of-hope
- https://www.mckinsey.com/~/media/McKinsey/Industries/Consumer Packaged Goods/Our Insights/Supply Chain 4 0 in consumer goods/Supply-Chain-4-0-in-consumer-goods-vf.pdf
- https://www.eiu.com/n/consumer-goods-retail/
- https://www.bain.com/insights/consumer-products-report-2024-resetting-the-growth-agenda/

-
https://kpmg.com/xx/en/home/insights/2023/12/supply-chain-trends-2024.html
-
https://deloitte.wsj.com/cmo/2024-consumer-products-outlook-turning-to-profitable-volume-34037ce4
-
https://link.springer.com/article/10.1007/s10479-023-05782-9
-
https://www.method.me/blog/efficient-supply-chain-management/
-
https://www.investopedia.com/ask/answers/042115/how-market-failure-corrected.asp
-
https://www.layerise.com/resources/blog/post/strategies-to-accelerate-digital-transformation-in-consumer-goods-companies
-
https://www.ascm.org/making-an-impact/research/top-supply-chain-trends-in-2024/
-
https://rsmus.com/insights/industries/consumer-goods/consumer-goods-outlook.html
-
https://www.forbes.com/sites/googlecloud/2021/06/08/investing-for-cpgs-digital-transformation-in-the-new-normal/
-
https://www2.deloitte.com/us/en/pages/operations/articles/digital-transformation-trends-for-cpg.html
-
https://www.bcg.com/publications/2024/how-consumer-goods-companies-win-in-turbulent-times
-
https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/winning-with-it-in-consumer-packaged-goods-seven-trends-transforming-the-role-of-the-cio
-
https://www.natlawreview.com/article/top-10-legal-issues-2024-consumer-products-industry
-
[https://www.ey.com/en_us/consumer-products-retail/how-embracing-technology-trends-can-

drive-leadership-in-the-next](https://www.ey.com/en_us/consumer-products-retail/how-embracing-technology-trends-can-drive-leadership-in-the-next)
-
https://www.salesforce.com/blog/consumer-goods-industry-trends/
-
https://www.eiu.com/n/wp-content/uploads/2023/10/Consumer-report-2023.pdf
-
https://www2.deloitte.com/us/en/events/industries-dbriefs-webcasts/2024/2024-retail-and-consumer-products-industry-outlooks.html
-
https://socialsci.libretexts.org/Bookshelves/Economics/Economics_(Boundless)/15:_Challenges_to_Efficient_Outcomes/15.1:_Sources_of_Inefficiency
-
https://www.netsuite.com/portal/resource/articles/inventory-management/supply-chain-efficiency.shtml
-
https://www.kadence.com/en-us/15-consumer-trends-to-watch-in-2024/
-
https://www.cxotoday.com/news-analysis/how-digital-tech-is-paving-the-way-for-fmcg-innovation/
-
https://www.mckinsey.com/industries/consumer-packaged-goods/our-insights/supply-chain-4-0-in-consumer-goods
-
https://www.investopedia.com/terms/m/marketfailure.asp
-
https://research.aimultiple.com/supply-chain-best-practices/
-
https://www.forbes.com/sites/paulnoble/2024/03/11/the-future-of-supply-chain-technology-a-shift-toward-intelligent-systems/
-
https://kpmg.com/xx/en/home/insights/2024/03/kpmg-global-tech-report-2023-consumer-and-retail-sector-insights.html
-
[https://www.researchgate.net/publication/371755619_FMCG's_Digital_Dilemma_The_Cons

equences_of_Insufficient_IT_Expertise_in_the_Fast-Moving_Consumer_Goods_Industry](https://www.researchgate.net/publication/371755619_FMCG's_Digital_Dilemma_The_Consequences_of_Insufficient_IT_Expertise_in_the_Fast-Moving_Consumer_Goods_Industry)
- https://socialsci.libretexts.org/Bookshelves/Economics/Economics_(Boundless)/7:_Market_Failure:_Externalities/7.1:_Introducing_Market_Failure
- https://www.ascm.org/globalassets/ascm_website_assets/docs/top-10-trends-report-2024.pdf
- https://www.foley.com/insights/publications/2023/12/trends-2024-consumer-packaged-goods-industry/
- https://www.dynamicyield.com/wp-content/uploads/2024/03/Consumer_goods_industry_trends_2024_Report_V2.pdf
- https://kpmg.com/us/en/media/news/2024-kpmg-supply-chain-stability-index.html
- https://openstax.org/books/introduction-business/pages/12-4-using-supply-chain-management-to-increase-efficiency-and-customer-satisfaction
- https://www.mastercardservices.com/en/advisors/data-strategy-management/insights/consumer-goods-industry-trends-2024
- https://blogs.sw.siemens.com/consumer-products-retail/2024/03/20/digital-transformation-in-cpg-future-personalized-products/
- https://hyscaler.com/insights/ai-in-fmcg-11-strategies-increase-efficiency/
- https://www.deloitte.com/cbc/en/Industries/consumer/analysis/consumer-products-industry-outlook.html
- https://www.shipbob.com/blog/supply-chain-efficiency/

- https://www.netsuite.com/portal/resource/articles/erp/supply-chain-best-practices.shtml
- https://www.mckinsey.com/industries/consumer-packaged-goods/our-insights/a-new-model-for-the-consumer-goods-industry
- https://www.gartner.com/en/newsroom/press-releases/2024-03-20-gartner-identifies-top-trends-in-supply-chain-technology-for-2024
- https://hbr.org/2024/03/how-machine-learning-will-transform-supply-chain-management
- https://throughput.world/blog/supply-chain-efficiency/
- https://www.gartner.com/smarterwithgartner/6-strategies-for-a-more-resilient-supply-chain
- https://www.mckinsey.com/industries/consumer-packaged-goods/our-insights/digital-innovation-in-consumer-goods-manufacturing
- https://www.apqc.org/resource-library/resource-listing/2024-supply-chain-priorities-and-challenges-executive-summary
- https://www.forbes.com/sites/sap/2021/06/17/tackling-the-biggest-issues-facing-the-consumer-products-industry/