AI IN SUPPLY MANAGEMENT: THE USE OF AI-BASED LANGUAGE MODELS TO IMPROVE EFFICIENCY IN PROCUREMENT CONTRACT MANAGEMENT

Lappeenranta–Lahti University of Technology LUT

Master's Programme in Supply Management, Master's Thesis

2025

Otto Classen

Examiners: Professor Katrina Lintukangas

Associate Professor Mika Immonen

ABSTRACT

Lappeenranta–Lahti University of Technology LUT

LUT Business School

Business Administration

Otto Classen

AI in supply management: The use of AI-based Language Models to Improve Efficiency in Procurement Contract Management

Master's thesis

2025

73 pages, 8 figures, 9 tables, and 1 appendix

Examiners: Professor Katrina Lintukangas and Associate Professor Mika Immonen

Keywords: Artificial Intelligence, Procurement, Contract Management, Prompt engineering

This thesis examines how artificial intelligence can improve the efficiency of work related to procurement contracts. The research includes applications, challenges, and the potential of AI tools to automate and support contract-related tasks, such as clause comparison, risk detection, and document summarization. The thesis uses a mixed-methods approach, consisting of a survey, a case analysis of prompt types on actual contracts, and a comparative analysis of AI models.

The survey aimed to evaluate sourcing managers' familiarity with AI tools and their perception of efficiency gains and challenges. The survey got responses from sourcing managers in Finland, France, and Germany. The findings showed that awareness and training are limited, but the perceived potential for efficiency improvement is high. Open-ended questions highlighted the importance of prompt design.

The empirical testing compared two LLMs across five prompt types to assess how their structure affects performance. The tests showed significant differences in model performance and underscored that AI's usefulness in contract analysis depends heavily on the clarity and context of the prompt.

This thesis concludes that, by combining human oversight with effective prompting, AI-based language models can significantly improve efficiency in procurement. However, these benefits require user training, responsible data governance, and alignment between new technologies and organizational processes.

TIIVISTELMÄ

Lappeenrannan–Lahden teknillinen yliopisto LUT

LUT-kauppakorkeakoulu

Kauppatieteet

Otto Classen

Tekoäly hankintatoimien johtamisessa: Tekoälyn käyttö hankintasopimusten hallinnan tehokkuuden parantamisessa

Kauppatieteiden pro gradu -tutkielma

2025

73 sivua, 8 kuviota, 9 taulukkoa ja 1 liite

Tarkastajat: Professori Katrina Lintukangas ja Tutkijaopettaja Mika Immonen

Avainsanat: Tekoäly, Hankinta, Sopimustenhallinta, Promptaus

Tämä tutkimus tarkastelee, kuinka tekoäly voi parantaa hankintasopimuksiin liittyvän työn tehokkuutta. Tutkimus käsittelee tekoälyn käyttöä, haasteita ja mahdollisuuksia automatisoida ja tukea sopimuksiin liittyviä tehtäviä, kuten ehtojen vertailua, riskien tunnistamista ja asiakirjojen tiivistämistä. Työssä käytetään monimenetelmällistä lähestymistapaa, joka koostuu kyselystä, erilaisten promptien analyysista sekä tekoälymallien vertailuanalyysista.

Kyselyn tarkoituksena oli arvioida hankintapäälliköiden tekoälytyökalujen tuntemusta sekä heidän näkemyksiään sen tehokkuushyödyistä ja haasteista. Vastauksia saatiin hankintapäälliköiltä Suomesta, Ranskasta ja Saksasta. Tulokset osoittivat, että yleinen tietoisuus tekoälyn käytöstä ja koulutus sen käytöstä olivat rajallisia, mutta koettu potentiaali tehokkuuden parantamiseen on korkea. Avoimet kysymyksen korostivat etenkin promptien suunnittelun merkitystä.

Tutkimuksen empiirisessä osuudessa verrattiin kahta kielimallia viiden erilaisen promptin avulla, jotta voitiin arvioida, miten niiden rakenne vaikuttaa suorituskykyyn. Testit osoittivat merkittäviä eroja mallien välillä ja näyttivät, että tekoälyn hyödyllisyys on riippuvainen kielimallista, promptin selkeydestä ja kontekstista.

Työn lopputulokset osoittavat, että yhdistämällä hankintapääliköiden ammattitaidon ja tehokkaan promptauksen tekoäly voi merkittävästi parantaa hankintatoimen tehokkuutta. Näiden hyötyjen saavuttaminen edellyttää kuitenkin käyttäjien kouluttamista, vastuullista datahallintaa sekä uusien teknologien ja organisaatioprosessien yhteensovittamista.

Abbreviations

AI	Artificial Intelligence
ERP	Enterprise Resource Management
GDPR	General Data Protection Regulation
HCI	Human-Computer Interaction
IP	Intellectual Property
LLM	Large Language Model
ML	Machine Learning
MSA	Master Service Agreement
NDA	Non-Disclosure Agreement
NLP	Natural Language Processing
PQM	Prompt Quality Model
RAG	Retrieval-Augmented Generation
SRM	Supplier Relationship Management
WPA	Work Package Agreement

Declarations

The author of this thesis, Otto Classen, used the following AI tools during the preparation of the thesis:

1. ChatGPT
    a. Purpose of use: To support academic writing, structure development, and language refinement during the thesis process.
    b. Explanation of the use of the tool: ChatGPT was used to assist in improving the structure, clarity, and academic tone of the text. It was also used to summarize literature and generate preliminary outlines. The tool was not used to produce or analyze research data. All ideas, interpretations, and final formulations were written, reviewed, and verified by the author.

2. Grammarly
    a. Purpose of the use: To ensure grammar correctness and consistency of writing style.
    b. Explanation of the use of the tool: Grammarly was used to identify grammatical and stylistic errors to improve readability and coherence throughout the text. All suggested edits were manually reviewed and accepted or rejected by the author to maintain full control over the final wording and meaning.

Responsibility

The author, Otto Classen, takes full responsibility for the content of this thesis and has reviewed and edited the content generated by the possible use of AI tools.

Table of contents

Appendices

1 Introduction

The increasing adoption of artificial intelligence (AI) across industries is reshaping the business landscape as we know it. AI can increase efficiency, contribute to value creation in businesses, and perform analysis based on the data it is given. However, to use AI and AI-based language models as reliable tools, companies face challenges. These challenges can range from technical issues, such as the data the language model has access to, to human-related barriers caused by a lack of understanding on how it should be implemented. This can ultimately lead to gains that do not match the expectations of this digital transformation. (Fenwick et. al. 2024)

This thesis aims to shed light on the correct usage and testing of large language models for specific use cases in procurement, enabling their consistent and reliable application. It aims to identify ways to enhance efficiency, mitigate risks, and enhance contract visibility through targeted testing and analysis. Contract management is often time-consuming, involving document review, compliance checks, and negotiation activities that require both accuracy and contextual understanding (Zou, et al. 2019). AI can assist in this by reducing and automating manual tasks, thereby speeding up processes while reducing the likelihood of human errors (Srivastava et al. 2024).

Procurement contracts often consist of multiple interlinked documents, including annexes and appendices that reference one another. This creates a risk of inconsistencies and can also lead to important clauses being formatted incorrectly or missing. AI-powered tools can help mitigate these risks by efficiently comparing contract versions and identifying discrepancies between templates and final agreements. Such capabilities not only enhance accuracy but also provide procurement professionals with a more reliable foundation for negotiation and compliance monitoring.

1.1 Key concepts

Digital transformation in procurement refers to the integration of advanced digital technologies to improve procurement processes, making them more efficient. Digital transformation represents a shift from manual labour to automated and cloud-based systems even aided by AI technology. Digital transformation in procurement is a key step in modernizing contract management. (Gong & Ribiere 2021)

Contract management is the process of handling contracts related to procurement. Contracts serve as a legal framework for establishing relationships with suppliers. It involves negotiating contract terms and conditions, ensuring compliance, and formally documenting all modifications approved by all parties involved. (Zou, et al. 2019)

AI-based language models or natural language processing is a branch of artificial intelligence that enables computers to interpret and process language similar to humans. Unlike traditional computing, it does not rely on predefined rules and data processing. AI can also recognize patterns which allows for more accurate results that can also vary based on the context that has been provided. (Sabri et.al. 2025)

1.2 Research goal and questions

The case company for the study is a Finnish subsidiary of a large multinational company working in the public safety and security sector. The goal of this research is to identify areas where stakeholders within the company can utilize AI tools to enhance their efficiency and accuracy in contract management. Another important goal of the thesis is to increase the knowledge of procurement processes, contract management, AI tools, and general AI implementations for the researcher. The empirical part of the thesis will present the current state of AI tool usage by procurement managers in the case company. The current state is crucial for the analysis, as it will serve as the basis for selecting specific cases in the testing phase. The thesis is based on a main research question and three supporting sub-questions.

Main research question: How can AI-based language models be used in Supply Management?

This research question is central for understanding how new artificial intelligence technologies, such as large language models LLMs can drive up the efficiency, accuracy, and strategic value of supply management processes (Srivastava et al. 2024). As supply chains become more complex, the number of contracts and supplier data increases for procurement managers. Traditional methods often rely on manual analysis, which limits both speed and consistency in decision-making (Van Weele 2018).

Sub-question 1: How can AI-based language models be used to improve the efficiency of work related to procurement contracts?

Procurement contract management is a complex process that involves drafting, reviewing, negotiating, and monitoring agreements to ensure compliance and minimize risks (Zou, et al. 2019). AI-based language models, particularly those leveraging Natural Language Processing (NLP), have the potential to automate and streamline many of these tasks. By enhancing efficiency, AI can reduce the time spent on contract analysis, improve accuracy, and enable procurement professionals to focus on higher-value tasks. This research aims to explore the role of AI in contract management, identify key efficiency improvements, and assess the broader impact of AI adoption in procurement processes (Dilmegani, 2025).

Sub question 2: How can AI models help identify risks, inconsistencies, or non-compliance in procurement contracts?

Procurement contracts include complex legal language, specific regulations, and compliance regulations. Manually reviewing these details can be resource-intensive and prone to human errors (Surden, 2019). This sub-question investigates the extent to which AI models can assist procurement managers in identifying risk, inconsistencies, or non-compliance in procurement contracts. The aim is to assess how reliably AI models can perform risk identification tasks and what level of human supervision is required.

Sub-question 3: What to take into consideration when writing a prompt? Including prompt engineering and context engineering.

The effectiveness of AI-based language models depends significantly on how prompts are structured and contextualized. Writing an effective prompt requires clarity, specificity, and appropriate contextual framing to produce accurate and relevant responses (Liu et al. 2023). This sub-question explores how prompt engineering and context engineering techniques influence the performance of LLMs in procurement-related tasks. The goal is to identify best practices that enable users to communicate effectively with AI systems and to ensure consistent, high-quality outputs in contract management and other supply management applications.

1.3 Research method

This thesis follows a mixed-methods research approach to understand how AI can be used effectively in supply management, specifically in managing procurement contracts. The study is conducted as a case study, a research method that involves an in-depth examination of a particular event, method, or thing that is significant, critical, or interesting. The case study method often serves as an umbrella term including multiple methods and data types, which makes it applicable across both qualitative and quantitative research. The primary objective of a case study is to provide comprehensive insights into the specific processes of the chosen case, often integrating multiple data sources and research methods to enhance the depth and validity of findings (Yin, 2003, 1-5).

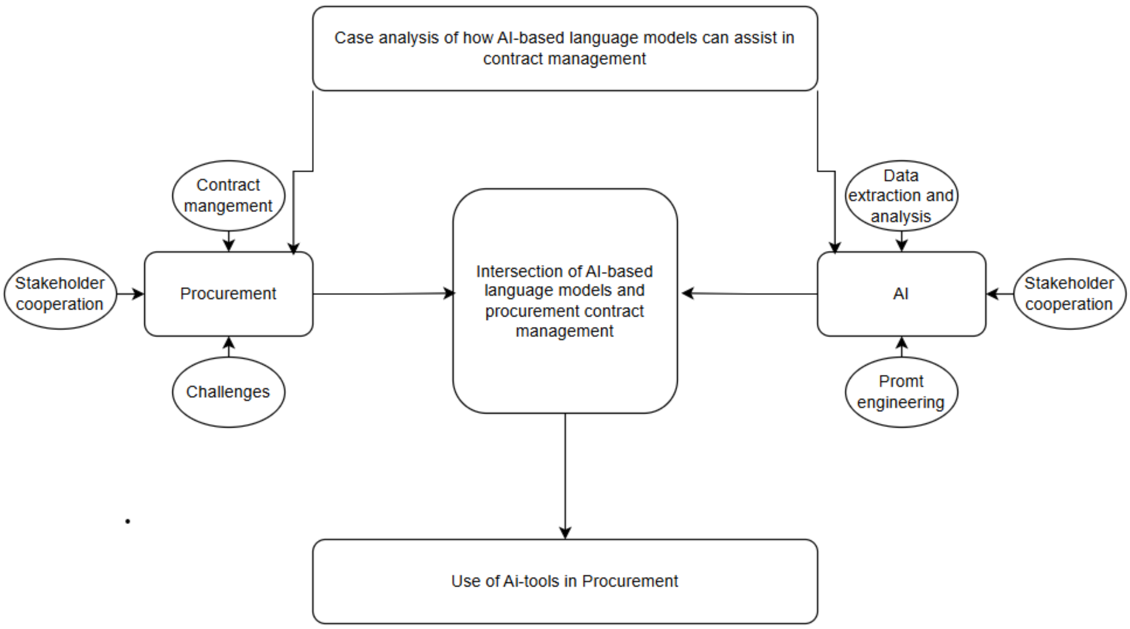
Another advantage of the case study method is its flexibility, as case studies can utilize different theoretical perspectives and methodological frameworks. This approach is particularly useful when research seeks a comprehensive understanding rather than merely statistical generalizations. Therefore, a case study can facilitate the uncovering of deeper insights into specific contexts, providing robust empirical grounding for theoretical development (Korhonen, 2009).

Case studies differ from other empirical research mainly on two distinguishing factors. The first factor is that the study is a contemporary phenomenon, and the other is that it has a real-life context (Wohlin, 2021). Both factors are found in this thesis.

1.4 Framework of the thesis

The conceptual framework (Figure 1) provides a broad overview of the research in the thesis. This thesis is positioned at the intersection of contract management in procurement and artificial intelligence-based language models, highlighting how these two domains interact to address practical challenges in sourcing environments. The framework illustrates how inefficiencies in traditional contract analysis, such as manual review, difficulties in identifying key terms in complex legal documents, and the risk of human error can be mitigated through the integration of AI tools that enhance information processing and decision-making. The specific issues addressed in this thesis include inefficiencies in time-consuming contract analysis, difficulties in extracting key terms from complex contracts, and human error. These are areas where AI-based language models can provide valuable support by clarifying terminology, summarizing clauses, and highlighting inconsistencies (Lai et al. 2023). These issues form the basis for the main research question.

Figure 1: Conceptual framework for the thesis



The main stakeholders in the study are sourcing managers who are responsible for contract creation, supplier evaluation, and negotiation activities. Sourcing managers operate at the core of procurement processes and must ensure that contractual terms align with both organizational objectives and supplier capabilities. These managers also collaborate with other internal stakeholders, such as project managers and legal teams, to align the purchase scope and ensure it is legally sound. Therefore, artificial intelligence language models can help bridge the gap in understanding the commercial and technical aspects of contracts. The efficiencies and accuracy gains in the framework aim to answer Sub-question 1, which explores how AI can support contract-related work.

This thesis is situated within the broader context of digital transformation. In digital transformation, organizations seek to improve automation and efficiency through the integration of digital technology to create value (McKinsey & Company, 2024). The adaptation of AI tools in procurement represents a concrete step towards this transformation. In this transformation, human AI collaboration becomes increasingly important because AI's task is to assist with data processing, while sourcing managers retain the final responsibility for risk assessment and contractual decisions (Amershi et al., 2019). This connection is linked to sub-question 2.

The framework emphasizes the importance of prompt engineering, which also includes context engineering. Prior research by Liu et al. (2023) demonstrates that well-structured prompts significantly improve the AI's response quality and consistency. This makes it a crucial part of the framework and links to sub-question 3, which investigates what must be considered when writing a prompt.

1.5 Structure of the thesis

The thesis is structured into several sections, including the theoretical foundations, methods, empirical analysis, and conclusions that each support the implementation of AI tools in procurement. The introduction presents the overall background and motivation for this study. It introduces the research goals, methodology, and conceptual framework of the thesis. It also explains the relevance of AI in procurement and defines the scope.

The theoretical part provides a comprehensive literature review of the theoretical foundations relevant to the thesis. It begins by discussing procurement and contract management, showcasing their role in supply management and organizational performance. The theory continues with the introduction of artificial intelligence, focusing on aspects such as natural language processing, human-computer interaction, and prompt / context engineering. The chapter also discusses the risks and opportunities associated with AI adoption in procurement, establishing the theoretical basis for the case study.

The research methodology and data collection chapters present the mixed-methods approach used in the thesis and explain why a case study method is appropriate for evaluating AI use in contract management. This section also describes the data collection methods, including surveys and empirical tests, as well as the reasons for the thesis's validity and reliability.

The main section of the thesis is formed by the case analysis. It begins with the survey analysis, which shows the current use of AI among sourcing managers in the case company. It then presents the case selection for testing of different prompt types and AI models. The results of this case study include both the performance of models and the qualitative differences between prompt structures. These results demonstrate how AI tools process contract information, identify risks, and support sourcing managers in their daily work

In the final chapters, the empirical results and connected to the theoretical basis of the thesis. The discussion interprets the findings in relation to the research questions and discusses both theoretical and practical implications. The chapter also addresses the limitations of the thesis and suggests areas for future research, particularly regarding broader testing and long-term effects of AI use.

2 Procurement and AI

This chapter provides an understanding of procurement and contract management in the context of business operations. The chapter begins with a definition of procurement and contract management. After this, it dives deeper into the importance of procurement contracts regarding risk mitigation and supplier relationship management. In addition, this chapter details how companies currently manage the entire lifecycle of procurement contracts, and it also discusses the challenges faced in contract management. This chapter is important because it creates the basis for understanding how AI-based solutions can improve contract management in the following chapters.

Procurement is a core function within supply chain management, responsible for sourcing and managing input from external suppliers to meet organizational needs. The role of procurement also becomes increasingly important as organizations increasingly rely on external resources to achieve operational efficiency and strategic goals. Van Weele (2018) defines purchasing as “the management of the company’s external resources in such a way that the supply of all goods, services, capabilities and knowledge which are necessary for running, maintaining and managing the company’s primary and support activities is secured at the most favorable conditions”. This definition underscores purchasing not merely as a transactional function but as a strategic activity aimed at securing value from suppliers and ensuring business continuity. In this view, purchasing extends beyond price negotiation to include relationship management, supplier evaluation, and risk mitigation.

Lysons and Farrington (2012) explain procurement further as an operational and demand-driven process in addition to its broad scope. Operational purchasing is a reactive and transactional function responding to internal organizational needs. Therefore, operational procurement primarily focuses on cost control, order accuracy, and timely delivery. The scope here is narrower, focusing primarily on acquiring goods and services.

The definition by Swink et al. (2020) aligns with Van Weele's, as it emphasizes that procurement is a key function in maintaining a company’s ability to control costs, ensure timely delivery, maintain quality, and build a resilient supplier base. The book separates procurement from purchasing by positioning procurement as a more strategic function with long-term goals in supplier relationship management and with more cross-functional alignment between finance and product development.

Procurement is commonly categorized into two types: direct procurement, which involves the purchase of materials directly used in the production of goods, and indirect procurement, which refers to goods and services that support day-to-day operations but are not part of the manufacturing process. Indirect goods include, for example, office supplies, IT equipment, maintenance and repair items, and various service-related purchases such as travel and marketing support (Kim & Shunk, 2004).

There is no distinct definition of procurement; however, the literature indicates that it is a multidimensional function with both operational and strategic roles within supply chain management. The “definitions” of procurement are collected in Table 1 below.

Table 1: Comparison of procurement definitions in literature (according to Van Weele, 2018; Lysons & Farrington, 2012; Monczka et al. 2014; Swink et al. 2020 and Joyce, B. 2006)

Author	Emphasis	Nature of purchasing	Scope
Van Weele (2018)	Strategic, resource based	Proactive, value-focused	Broad: goods, services, knowledge and supplier management
Lysons & Farrington (2012)	Operational, demand driven	Reactive, transactional	Moderate: goods and services
Monczka et al. (2014)	Functional, organizational	Role-based, process focused	Moderate: materials and services
Swink et al. (2020)	Operational, strategic	Proactive, decision driven	Broad: goods, services, knowledge and supplier management
Joyce, B. (2006)	Strategic and operational	Value-driven, cost-sensitive	Moderate: materials, parts, services

The paper by Joyce, B. (2006) emphasizes that purchasing is not only a transactional role. According to Joyce procurement has a direct impact on a company’s cost structure, product quality, and operational performance. Particularly, manufacturing costs can be heavily influenced by purchased components since they often comprise over 60% of total costs in manufacturing firms and over 90% in retail. The role of purchasing therefore includes supplier selection, negotiation, contract management, and make-or-buy analysis.

2.1 Procurement processes

The procurement function traditionally includes processes for obtaining materials and services. Beyond sourcing, procurement processes play an important role in determining a company’s profitability and competitive edge. Such activities include negotiating purchase prices, defining contractual terms and conditions, managing contracts, and monitoring purchases to confirm timely delivery and adherence to agreed quality standards (Van Weele, 2014).

Procurement functions can be categorized into strategic and operational activities. Operational purchasing typically includes supply-related tasks, whereas strategic purchasing involves longer-term decision-making processes such as supplier selection and contract management. Notably, strategic procurement explicitly excludes routine activities, such as everyday ordering, logistical coordination, and financial transactions (Sollish, 2011, 2).

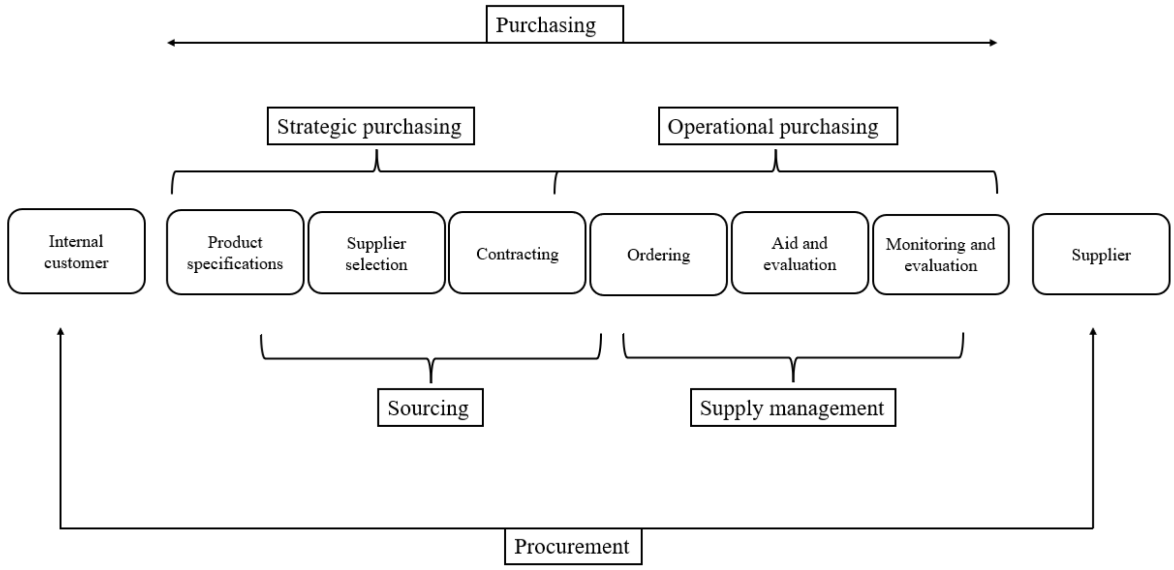


Figure 2: Procurement processes (Adapted from Van Weele. 2014, 9)

Figure 2 illustrates a purchasing process model highlighting essential procurement functions and their interrelationships. Van Weele (2014, 9–10) notes that this model specifically excludes inventory-related activities, such as inventory planning, scheduling, and stock management, which fall under the supply management category.

2.2 Supplier relationship management

Supplier Relationship Management (SRM) is a strategic approach to the management of supplier interactions. The goal of SRM is to maximize the value of buyer-supplier relationships. Effective supplier relationship management is especially important during times of uncertainty and supply chain disruptions, such as the ones seen during the COVID-19 pandemic. During those times, companies have to use an increasing amount of resources to ensure deliveries and the supply of materials (Wasthy et. al. 2022). In a study, Tarigan et al. (2019) stated that to increase profitability, companies should focus mainly on their core business and not produce their own source of supply. This should be achieved through a network of trusted suppliers that can deliver high-quality materials at a competitive cost.

Therefore, SRM means going beyond transactional buying to creating long-term, mutually beneficial partnerships between companies and suppliers (Irshadullah, 2024).

In previous studies, Supplier Relationship Management has often been described as a multi-phase process consisting of various activities (Lambert & Schwieterman, 2012; Park et al., 2010). Procurement can be divided into strategic and operational functions as visualized in Figure 2, "Procurement processes". The division can be done for SRM. Lambert and Schwieterman emphasize in their article that in strategic SRM, suppliers are differentiated based on their significance and contributions. Managers then tailor relationship strategies for different suppliers. For example, dedicating more resources and collaborative efforts to high-value or strategic suppliers. This high-level planning aligns supplier capabilities and development with the company's long-term goals. At the operational level, SRM involves the day-to-day management of supplier performance. Operational tasks include continuous communication, performance evaluation, and joint problem solving.

According to Park et al. (2010), SRM can be broken down into four major phases, as seen in Figure 3: purchasing strategy, supplier selection, collaboration, and supplier evaluation and development. Their framework illustrates this as a structured process involving the formulation of purchasing strategies, the choice of suitable suppliers, collaborative interaction, supplier assessment and development, and continuous improvement. Through these steps, companies can ensure that suppliers align with corporate strategy, while assessment results guide decisions on supply volumes, supplier support, and further development.

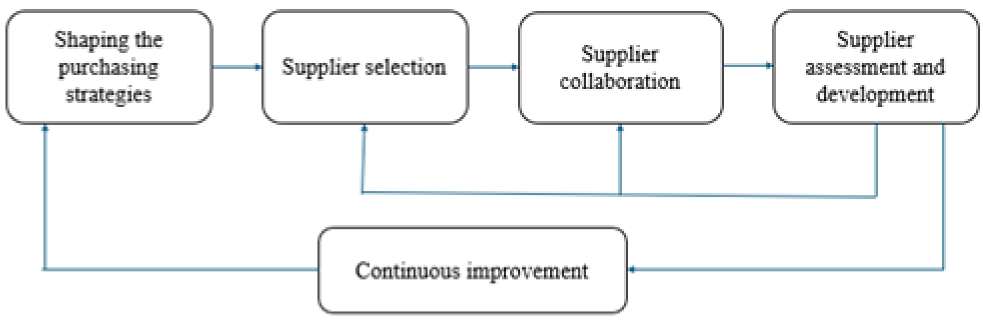


Figure 3: Framework for Supplier Relationship Management (Park et al. 2010)

A key element of SRM is also the use of digital technologies to strengthen supplier management. Solutions such as supplier portals, advanced analytics, and automated communication platforms offer real-time insights into supplier performance. This simplifies procurement activities and supports informed decision-making. Implementing new technologies into SRM enables companies to detect risks early, ensure compliance, and increase transparency across their supplier base.

Through the research in this thesis, it has become evident that Supplier Relationship Management is a key element for achieving operational efficiency. It is, however, also increasingly recognized as a way to increase innovation and to reduce supply risks. Through SRM, companies can establish a foundation for collaboration and trust that extends beyond a transactional relationship.

Close supplier relationships offer opportunities for knowledge sharing and joint product development. Wieland and Ivens (2025) highlight in their study that effective SRM practices encourage suppliers to contribute expertise and technological capabilities, which support innovation. Research also shows that collaborative partnerships lower barriers to experimentation because trust reduces the perceived risk associated with innovation (Reynolds, 2024).

Another crucial role of SRM is its contribution to supply chain risk management. Long-term supplier relationships enable companies to identify early warning signs of disruptions, compliance issues, or financial instability, thereby functioning as an early detection system for supply risks (Santos et al., 2025). Furthermore, SRM enables collaborative risk mitigation, allowing buyers and suppliers to coordinate production schedules and share responsibilities for inventory buffers. This reduces vulnerability to external events (Mwesiumo et al., 2021).

2.3 Procurement contract management

Contracts are legally enforceable agreements between two or more parties. Contracts define the obligations of the parties involved. Simply put, contracts specify the deliverables for the seller and the corresponding payment for the purchaser upon delivery of the goods or services. Contracts should therefore always include the constraints under which the goods or services should be delivered, contract administration such as payment terms and change control, acceptance criteria for deliverables, and damages for non-performance. (APM, 2017 17-18)

According to the Association for Project Management (2017), the delivery phase begins once a contract is awarded to a supplier. In the delivery phase the supplier must deliver the agreed solution in line with the agreed provisions in the contract. This chapter of the thesis will go over the individual phases (Table 2) of contract management throughout the contract life cycle.

Table 2. Contract management phases (APM, 2017, 114-126)

Phase (APM 2017)	**Description**	**Contract management focus**
1. Initiation	Establish relationships, review contract, and clarify roles and responsibilities	Leadership, communication with stakeholders
2. Planning and definition	Confirm deliverables, dependencies, schedules and risk controls	Alignment with the supplier and internal stakeholders
3. Implementation	Design-build-deliver cycle, progress reviews	Monitoring
4. Risk management	Manage risks throughout delivery, including supplier risk	Proactive identification and mitigation
5. Change control	Evaluate and approve any contractual changes	Flexibility, documentation, cost/time control
6. Acceptance and closure	Certify completion, resolve defects, finalize payments	Verification and documentation

In the initiation phase of contract management, the parties align the contract's basic purpose and structure. This includes verifying assumptions, clarifying governance roles and responsibilities, and ensuring that contract terms are consistent. In the worst-case scenario, inconsistencies could invalidate the contract, but most likely they will cause delays and extra costs for either or both parties. During the planning and definition phase, deliverables and milestones are defined. This ensures that all parties agree on what is being delivered and how success will be measured. Edkins et al. (2013) highlight that early definition and stakeholder alignment significantly improve predictability and minimize later-stage renegotiations. These conclude the pre-implementation phases of contract management (APM, 2017, 114-126)

The implementation phase of contract management is the phase of execution and delivery. Implementation covers the actual delivery of goods or services and continuous progress tracking. In this phase, the responsible manager ensures that the supplier's performance aligns with the contract requirements. Another important task is to ensure that the potential deviations are corrected. This phase is essential because it directly determines whether the objectives in the contract are achieved. If monitoring during the implementation phase is not done, the work in the pre-implementation phases is effectively lost. (APM, 2017, 114-126) Transparent coordination with internal stakeholders and the supplier has a significant impact on delivery performance, as it has been found that, combined with clear contractual specifications, it can reduce opportunistic supplier behavior. (Klijn et al. 2024)

Harju et al. (2024) show that systematic risk management practices improve supplier performance and reduce disputes by maintaining continuous oversight. The risk management phase is vital throughout the contract life cycle. During the delivery, expected and emerging risks are identified and mitigated. Monitoring ensures that deliveries fulfill contractual and regulatory requirements.

The fifth phase of change control involves identifying, evaluating, and, if necessary, implementing modifications to a contract's scope, cost, schedule, or other objectives. According to Garrett & Rendon (2012), change control should follow a formalized and traceable process. For this, they propose a framework consisting of the following steps: (1) identification of the potential change, (2) documentation, (3) evaluation of impact, (4) negotiation, (5) approval, and (6) implementation and monitoring. By evaluating changes through this framework, managers can update contracts without losing the thought behind the original version. Garrett and Rendon highlight that undocumented changes are among the leading causes of contract disputes, cost escalation, and performance delays. The closure phase finalizes the contract by validating deliverables and making final payments.

2.4 AI in Procurement

This chapter is about the risks and opportunities associated with the use of AI in contract management. The first section of the chapter explores the opportunities, including efficiency improvements, cost savings, and risk mitigation. The risks of AI are addressed in the second section of the chapter. Those include ethical concerns, AI bias, and legal considerations which may lead to unwanted results. By evaluating both risks and opportunities, this chapter helps clarify the most effective approach to implementing AI in managing procurement contracts.

AI in procurement refers to using technologies like machine learning and natural language processing to automate, enhance, and optimize procurement tasks. The potential of AI in procurement is widely recognized, and new pilots and platforms are emerging regularly (Guida et al. 2023). The new trend of using AI for procurement purposes is also evident in the fact that the weekly use of generative AI tools in procurement rose sharply in 2024, with 94% of procurement executives now using generative AI at least once per week. Moreover, 80% of Chief Procurement Officers (CPOs) plan to deploy generative AI in some capacity in the next three years. (Idelson, 2025)

2.5 Opportunities

AI can increase efficiency, especially in information-intensive tasks, by automating data processing, pattern recognition, and feedback generation. Celik (2023) explains that AI can “capture data from the environment, process it, and produce environment-specific outputs.” This enables companies to handle large datasets faster. AI can therefore also be used in procurement to track supplier performance, highlight contractual deviations and, for example, generate alerts for upcoming renewals. As in other research, Celik also found that one of the main benefits of AI integration is that it increases the time managers can spend on strategic decision-making and value creation. AI can therefore contribute to both operational efficiency and improved governance in contract management.

AI technologies streamline procurement workflows, leading to faster cycle times and lower operational costs. By automating labor-intensive tasks like invoice processing or contract review, AI reduces manual effort and errors, resulting in direct efficiency gains. A McKinsey analysis estimates that effective use of data and AI can increase the pipeline of procurement value initiatives by up to 200%, indicating significantly more savings projects and opportunities identified than before. (Mittal et al. 2024)

AI improves an organization’s ability to manage risks in the supply base and ensure compliance with policies. AI tools can continuously monitor supplier data and external risk indicators to flag potential issues. Potential issues include a supplier being in financial trouble, a geopolitical event that could disrupt supply lines, or the detection of fraud in transactions. Being proactive in risk identification can help companies avoid or mitigate disruptions with their suppliers. For example, an AI system can qualify alternative suppliers in advance, reducing the time spent manually evaluating them. (Dilmegani, 2025)

Based on Dilmeganis' (2025) guide, AI systems are well equipped for checking procure-to-pay transactions against rules. This reduces the chance of compliance failures such as over-budget spending or violation of contract terms. Overall, AI provides greater transparency and control, making sure that procurement activities follow best practices and that errors or non-compliant actions are quickly addressed. This strengthens governance and reduces losses from unforeseen issues.

2.6 Artificial intelligence risks

As artificial intelligence becomes increasingly embedded in procurement processes, the importance of ethical considerations grows. While AI promises improvements in efficiency, decision-making, and supplier collaboration, it also introduces new risks related to bias, transparency, accountability, and data privacy. (Guida et al., 2023; Osmonbekov et al., 2024).

The "Black-Box" problem is linked to the ethical use of AI. At its core, the problem is that many advanced AI models produce outputs that can be difficult to understand. Especially in procurement, this raises the question of explainability and accountability. For example, if an AI system were to flag contract terms as risky, the users need to understand why. As the user becomes more comfortable with the AI, there is a risk that they will blindly follow its suggestions or fail to account for the risks. (Guidotti et al., 2018)

The increasing integration of artificial intelligence into fields such as procurement promises significant efficiencies and enhanced decision-making. However, it also introduces critical challenges, notably the risk of AI bias. AI bias occurs when the outputs of an AI system are influenced by biased data or flawed algorithms, leading to misleading and potentially harmful results. (Holdsworth, 2023)

In addition to the AI system being prone to having biases, it also creates user-related biases. When working with AI, managers might subconsciously favor AI suggestions that align with their pre-existing biases or experiences, while dismissing those that challenge their views. As seen in the study by Ha et al. (2024), in which users relied more heavily on AI during difficult tasks, sourcing managers might place too much trust in AI outputs for complex contracts or when under pressure, thereby forgoing sufficient critical review. This can lead to the unwanted use of biased AI recommendations.

The use of AI in procurement introduces significant data privacy risks, especially when models process personal or sensitive data. Kamyshanskaya (2021) emphasizes that companies can face legal liability not only for their own AI implementations but also for those used by third-party vendors. The primary concerns regarding data privacy are compliance with regulations such as the GDPR, which mandate transparency in the collection, processing, and sharing of data.

2.7 Existing AI tools for procurement

The use of AI tools in procurement has gained a lot of attention in recent years. One of the most used applications of AI in procurement is in contract analysis and management. AI-powered tools can help review, manage, and analyze contracts more efficiently. This can significantly reduce the time spent on manual reviews and minimize the risk of errors.

Existing AI tools include Kira by Litera, which is used for contract analysis. It uses machine learning and natural language processing to identify and extract key clauses from legal documents, making them easier to analyze and compare. Kira's AI is trained to recognize a wide range of contract clauses, including payment terms, delivery schedules, and intellectual property rights, thereby removing the need for user prompting. Litera claims that by automating the extraction of critical information, it reduces the time required for contract review and increases the accuracy of analyses (Litera n.d.)

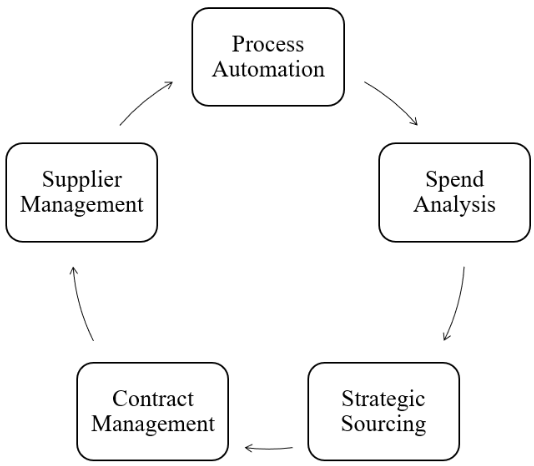


Figure 4: Use cases for artificial intelligence in procurement. (Gerber 2025)

Modern AI tools can support multiple phases of the procurement cycle (Figure 4). According to Gerber (2025), leading AI procurement platforms include SAP Ariba, Ivalua, and Suplari, which all integrate analytics, automation, and predictive insights into all procurement processes. Spend analysis systems can identify cost-saving opportunities and detect irregularities in purchasing patterns, whereas contract management and supplier management tools help by using AI tools to extract clauses and monitor supplier performance.

3 Artificial Intelligence technology

This chapter aims to give an understanding of how AI tools work. It begins with the introduction to AI-driven analysis, providing a base to how processes are automated through AI. Next, it examines the role of natural language processing in understanding and contract language, which helps the user in the extraction of data. This chapter also introduces prompt engineering and context engineering as key concepts in interacting with AI-based language models to generate precise answers with the correct context. This literature review provides a foundation for understanding AI technology.

3.1 Natural language processing

Natural language processing (NLP) is about bridging the gap between human communication and machine understanding. Computers are very effective at processing structured data. However, natural language, written by humans, is often more open to interpretation and context-dependent, posing significant challenges for machines. According to Jurafsky & Martin (2024), early NLP systems relied on rule-based methods, which limited their use because the rules for language structures had to be manually entered. Machine learning, particularly deep learning models, has enabled NLP to learn language patterns from large datasets. This has significantly improved their ability to process human-written language.

Figure 5 below shows the relationship between artificial intelligence, machine learning, and natural language processing, as well as their key components. It shows how NLP operates as a subfield of AI and is therefore connected to machine learning methods that enable the automation of language understanding

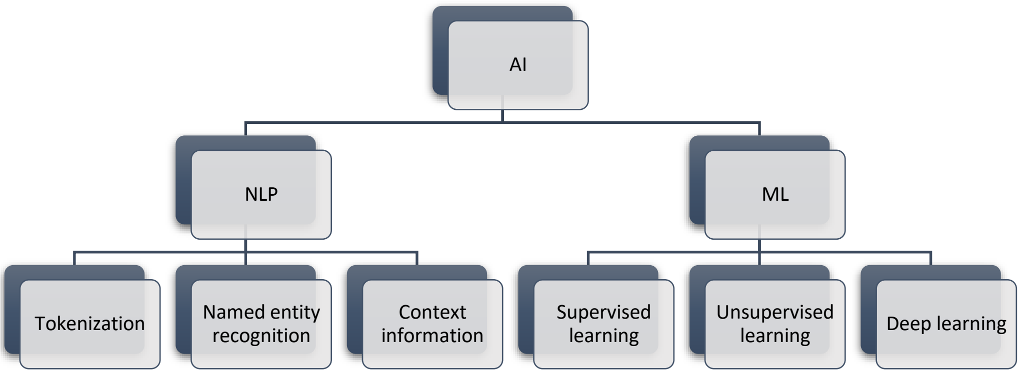


Figure 5: NLP and machine learning. Bista, N. n.d. (modified)

Key components of NLP include tokenization, named entity recognition, and context learning. Tokenization is a fundamental step in NLP that creates the basis for analysis by splitting text into smaller parts known as tokens. Named entity recognition is used to identify and classify entities within a text. By doing this, the model can better extract valuable information from the provided text and better understand its context. Context information involves looking at surrounding words, sentences, or the overall discussion to understand the true meaning of a word or phrase. It helps us interpret what came before and what follows, giving us a clearer picture of the intended message. This is particularly important since the context of words can change the entire message. (Kang et al. 2020)

3.2 AI learning types

Machine learning is a division of AI that involves using data to train models that can make predictions without explicit programming for each decision-making step. The key difference between machine learning and traditional programming lies in the fact that, while traditional software is managed by predefined rules and logic designed by human developers, machine learning is guided by data and algorithms. Machine learning systems evolve by learning from data. As more data is fed into the machine, its model improves, enhancing its ability to predict information based on the given inputs. (Merilehto, 2018, 27-67)

Machine learning algorithms can be classified into different types based on the tasks they perform, such as supervised, unsupervised, and reinforcement learning. In supervised learning, models are trained on labeled datasets, where the input data is paired with the correct output. The algorithm learns by analyzing the given examples, allowing it to predict results for new data. For example, when e-mail systems detect spam, they classify e-mails as either "spam" or "not spam". Once the model has been trained, it can classify new e-mails automatically (Kananen & Puolitaival, 2019, 45–55; Kolosova & Berestizhevsky, 2021, 9-10).

In unsupervised learning, the model tries to identify patterns and groupings in data without predefined labels (Kananen & Puolitaival, 2019, 45–55). Kolosova and Berestizhevsky (2021, 9-10) explain that unsupervised learning algorithms are designed to find structures in data that are not obvious. This can include tasks such as grouping similar contract types or finding purchasing patterns. Kristensen (2021, 71) explains unsupervised learning through an example where an image recognition system is given ten pictures of dogs. The system is then asked to filter the dogs based on breed. The system has no prior knowledge of dog breeds or even the fact that they are dogs, but through finding similarities, it is able to categorize them. In terms of procurement and contract management, unsupervised learning can be used to collect contracts with a similar risk profile, even if the contracts are not labeled based on risk.

Reinforcement learning, on the other hand, involves training a model through trial and error, with the model receiving rewards or penalties based on its actions, making it useful for decision-making tasks in dynamic environments (Schwartz, 2014, 13-14). In reinforced learning, an agent interacts with the environment and learns to take actions that maximize the rewards over time. Unlike in supervised learning, reinforcement learning does not rely on labeled data. It discovers optimal strategies by continuously interacting and adapting to the environment. In practice, reinforcement learning has been used successfully in logistics optimization and inventory management. (Nandy & Biswas, 2018, 5-22)

Deep learning is based on neural networks with many layers to process and learn from large amounts of data. These networks are designed to mimic the way the human brain processes information, making them particularly effective for tasks such as image and speech recognition, natural language processing, and autonomous systems. (LeCun, Bengio, & Hinton, 2015)

3.3 Human-computer interaction

Human-computer interaction (HCI) is a research area that examines how people interact with digital technologies. It aims to improve usability, accessibility, and overall experience of using computer systems (Dix et al. 2004). Norman (2013, 22) adds to that by explaining how the interaction design aims to increase the user's understanding of what can be done, what is happening, and what has just happened. By answering these questions, models become enjoyable to use. Further, Norman emphasizes the role of immediate feedback in HCI. Delays and unusable feedback can lead to frustration and inefficiency which does not motivate the user to keep using the model. Equally important as feedback is discoverability. Discoverability allows the user to understand what actions are possible and how to perform them. In LLMs this translates to the type of prompt the user uses and its clear context. (Norman, 2013)

In the context of artificial intelligence, HCI has evolved to understand how people, and AI s systems can collaborate effectively. In a study by Amershi et al. (2019), this is described as human-AI interaction. Their perspective emphasizes that the use of AI is not supposed to be a replacement for human decision-making but rather as partners that can mimic human capabilities. That is why efficient correction feedback loops are so important because being able to refine results without starting from the beginning resembles human interaction.

Amishi et al. (2019) argue that traditional human-computer interaction principles are no longer sufficient for AI systems, as they can produce inconsistent or unexpected results. To address these challenges they developed 18 design guidelines for effective human-AI interaction. The guidelines are categorized into four categories: Initially, during interaction, when wrong, and over time. The initial category emphasizes that system capabilities and limitations should be made clear to the user, and the importance of contextual information. During the interaction, the system's main goal is to be contextually relevant and to be objective in reasoning, meaning not to reinforce any stereotypes or biases. As stated before, feedback is crucial to user satisfaction; therefore, systems must be able to understand when the user feels like it is wrong and self-correct accordingly. Over time, in order to become a partner to the user, the system must remember recent interactions and learn from user behavior.

3.4 Prompt engineering

Prompt engineering is the practice of creating inputs that steer AI models to produce contextually appropriate and accurate outputs. Unlike traditional programming, this process relies on natural-language formulations, making it accessible to non-programmers while requiring a deep understanding of how LLMs interpret language. Korzynski et al. (2023) emphasize that prompt engineering is a new digital competence that integrates technical proficiency and human-centric creativity.

Liu et al. (2023) classify prompt formats into four primary categories, each with different strengths. The descriptive prompt is a general language input that leaves the interpretation open for AI. An example of this would be to ask the AI to summarize a given contract. This approach may work well if you are generally familiar with the context, as it creates a risk that AI could miss crucial information. The next step from descriptive prompts is creating instructional prompts, which are used to define explicit tasks. Instructional prompts create the opportunity to, for example, compare clauses in different contracts with prompts such as "Compare the liability clauses in these two NDAs." Another way to add more depth to prompts is to create role-based prompts, which involve framing the AI as a specific persona, such as instructing the AI to be a legal advisor when asking about contract clauses. With this additional context, the AI is able to tailor its responses according to their role. Step-by-step prompts explicitly request a structured breakdown of the process in the AI's response. For example, instructions such as "First summarize, then compare, then highlight risks" provide a clear, sequential framework that directs the AI to address each aspect of the task systematically. The step-by-step process is particularly beneficial in tasks that involve complex analysis such as contract comparison. By breaking down the task into steps, the user can also intervene if he or she notices inconsistencies and guide the AI back to the correct path.

To analyze and improve the quality of prompts, Koch et al. (2025) propose the Valisory Prompt Quality Model (PQM). The PQM (Figure 6) is a multidimensional framework that provides a structured method for optimizing how users communicate with AI. According to Koch et al. (2025), assessing prompt quality is a combination of style, word choice, and aligning the prompt's intent and tone with the AI’s mechanisms. The primary dimensions of the PQM model are: Relevance, Creativity and Originality, Linguistic Quality and Style, and Target Audience Appropriateness. Each of the dimensions includes three measurable aspects.

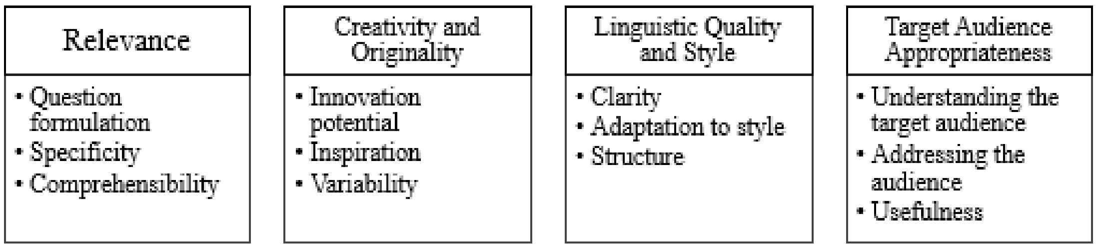


Figure 6: PQM model (Koch et al. 2025)

The dimension of relevance assesses the fundamental wording of the prompt and its appropriateness to the task's objective. This includes evaluating how the prompt is formulated, its specificity, and comprehensibility. The specificity of a prompt refers to how clearly its goals and expectations are defined, enabling AI to produce relevant answers. The comprehensibility aspect refers to how well the prompt is written, allowing the language model to understand what is expected of it. The dimension of creativity and originality evaluates the degree to which the variability of responses is constrained. Innovation potential measures whether the prompt promotes AI to create creative solutions that add unexpected value to the user. The dimensions of linguistic quality and style measure the clarity, style, and structure of the prompt. When formulating prompts, it is important to be as precise as possible. This means, for example, defining possible double meanings, especially if they can be used in a similar context. The target audience appropriateness dimension evaluates how effectively the prompt meets the expectations and requirements of the user. It begins by assessing whether the prompt clearly conveys to the language model who the target audience is, including their level of knowledge, interests, and informational needs. Unlike the linguistic quality dimension, which focuses on the wording of the prompt itself, this criterion emphasizes the inclusion of stylistic and contextual guidance aimed at a specific audience. (Koch et al. 2025, 24-27)

3.5 Context engineering

Context engineering is seen as the next step from prompt engineering. Prompt engineering involves the techniques used to create instructions for large language models in order to achieve effective results. In contrast, context engineering focuses on managing and refining the information that the model processes. This includes not only the prompt itself but also any additional contextual data that influences how the model interprets and generates its responses. (Rajasekaran et al. 2025)

Context engineering makes adaptive use of large language models possible by shaping the environment in which they operate. AI models can adjust to new tasks without needing re-learning through in-context learning. This adaptability is, however, heavily dependent on the context window size, which determines the amount of reference material the model can process simultaneously. While context engineering increases flexibility and efficiency, it does not fully resolve the reliability challenges of LLMs. These models still face issues such as hallucination and unfaithfulness to the provided information.

In large language models, the context is defined by several key components shown in Figure 7 below. The instructions serve as the foundation for the interaction. They determine the model's overall behavior, tone, and limits. These can include, for example, style guidelines or specific rules that the model must follow. It is similar to the user prompt, which represents the immediate prompt given by the user. The short-term memory refers to the current conversation history. Short-term information is used to provide continuity, allowing the user to have a coherent conversation with the model. Long-term memory extends beyond immediate information. Through long-term memory, language models can use data gathered over multiple sessions. (Philschmid, 2025)

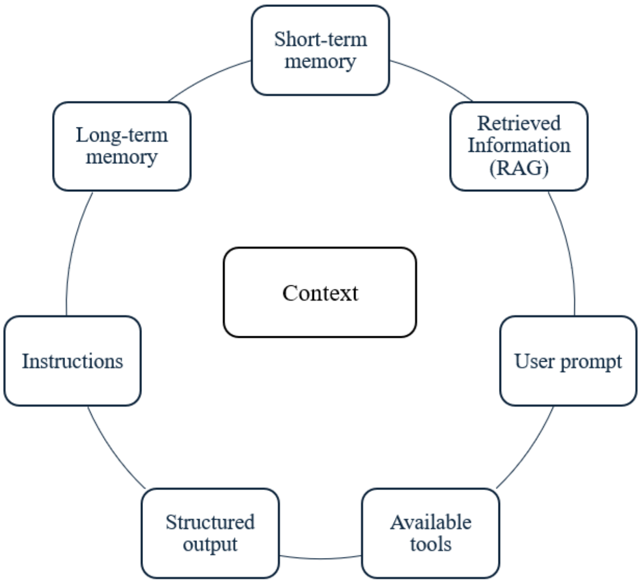


Figure 7: Context engineering (Philschmid, 2025)

Retrieved information (RAG) is used to increase factual accuracy by retrieving data from external sources such as documents and databases (Gao et al, 2023). The available tools are defined by the external functions the model can invoke, such as checking inventory from an ERP system, writing ready-to-send emails, or creating spreadsheets. The structured output specifies the response format that is wanted. This is linked to the available tools as they are the limiting factor to what the model can produce. (Philschmid, 2025)

3.6 Synthesis: Procurement and AI Applications

This chapter of the thesis combines the theoretical perspectives on procurement and artificial intelligence discussed in previous chapters. The perspectives are combined to yield a theoretical conclusion: when combined, these two aspects can enhance procurement contract management and the use of artificial intelligence in supply management fits in the broader trend of digital transformation (Gong & Ribiere, 2021; McKinsey & Company, 2024). In procurement, activities such as contract management, supplier relationship management, and risk control are information-sensitive and require extensive data analysis (Van Weele, 2018;

Monczka et al., 2020). These characteristics make them particularly suitable for implementing AI for automation, pattern recognition, and contextual reasoning (Balkan & Akyüz, 2025).

Through natural language processing, AI systems can analyze contextual data from procurement contracts. Machine learning enables models to detect patterns, identify differences between contracts, and classify them based on, for example, risk. Natural language processing further supports the use of AI tools for procurement tasks, as it can understand the meaning of legal text, enabling it to summarize and translate efficiently. (Jurafsky & Martin, 2024)

The theory also shows that the introduction of artificial intelligence tools into procurement is not purely technical. The role of human-computer interaction and prompt engineering is critical in making sure that users can efficiently communicate with LLM models, as explained by Norman (2013). In the procurement context, this means interfaces must be intuitive. Where HCI focuses on what the system must do for the user, prompt engineering and context engineering bridge the gap between human reasoning and machine processing. The quality of prompts and context defines how well the user can steer the model in the correct direction and how exact the responses are. Prompts engineering focuses on how instructions are phrased, while context engineering manages the surrounding information that gives the prompt meaning. (Liu et al., 2023; Koch et al., 2025)

To conclude, the theoretical part of this thesis shows that AI-based language models can serve as assistants within procurement, enhancing human decision-making rather than replacing it. This theoretical framework serves as the foundation for empirical testing and analysis of AI performance and prompt structures for procurement work.

4 Research methodology and data collection

This chapter discusses the chosen methodology, data collection, and data analysis. This part of the thesis also explains how testing for AI use in procurement contracts is conducted. This thesis uses a mixed-methods research methodology. The choice of a mixed approach stems from the nature of the research and the study's objective, which is to prioritize depth in understanding how AI can assist stakeholders. According to Creswell and Plano Clark (2018, 1-29), mixed-method research is suitable when the goal is to combine testing results with contextual understanding.

4.1 Research Process

The study is organized into four phases: First, the survey phase; Second, case selection; Third, testing; and finally, Fourth, analysis of the test results. The first phase of the process, visualized in Figure 8, involved surveying procurement professionals. The purpose of the survey was to understand current familiarity with AI tools for contract sourcing and management, and to identify potential areas where AI could be applied. In addition, the survey was conducted to act as a basis for the case selection by highlighting the most relevant use cases and contract-related challenges where AI could deliver efficiency improvements. Choosing the case based on user input ensured that the following stages of the research remained closely aligned with real-world procurement needs.

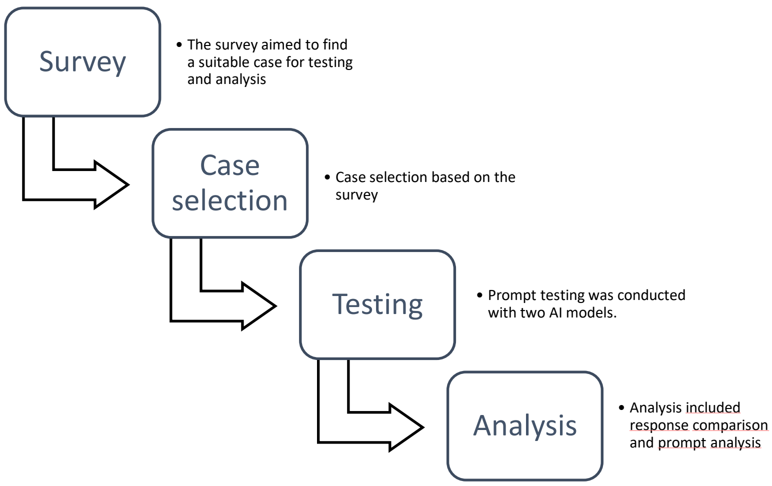


Figure 8: Research process description

The survey received 11 responses from sourcing managers based in Finland, France and Germany, with varied experience levels. As seen in Table 3 the levels of experience varied from 4 to over 21 years, with the majority having over 10 years of experience in the field of procurement. This indicates that most of the managers have worked a lot with contract management, supplier negotiations, and other procurement activities. This extensive experience makes their insights valuable for evaluating AI's potential. The respondents represented multiple procurement functions, including industrial sourcing, project sourcing, and general procurement, which reflects a wide range of needs and responsibilities.

Table 3: Background on survey respondents

Sourcing area	**Years of sourcing experience**	**Familiarity with AI (based on survey analysis)**
Industrial sourcing	6 years	Used Model B, dissatisfied
General procurement	15 years	Trained user, confident, efficient
Industrial souring	5 years	Familiar with AI, structured, cautious
General procurement	8 years	Experienced user, prompt-savvy, no training
Industrial sourcing	> 15 years	Trained, limited use, struggles with prompts
Industrial sourcing	21 years	No experience, theoretical understanding
Project sourcing	17 years	No training, cautious, open to use
Project sourcing	4 years	No AI or contract experience
Industrial sourcing	10 years	Uses AI for various tasks related to sourcing
Project sourcing	4 years	No AI tool experience
Industrial sourcing	>10 years	No use, strong theoretical grasp

The survey (Appendix 1) included a variety of multiple-choice and open-ended questions, tailored to support the sub- and main research questions of the thesis. In total, the survey consisted of four sections and a total of 18 questions. The questions started with basic background information about the sourcing managers after which they were asked about their current use of AI tools. The third section of the survey was about the sourcing managers perspectives on AI's potential. The questions in this section had agree/disagree questions about the use of AI and an open question where the sourcing manager had to formulate a prompt for a predefined task. After the survey was created, it was shared within the target company. The survey was only shared with members of the procurement team, as the scope of the thesis was to evaluate AI use in procurement. The interview was open for 10 days, after which is was closed for analysis and case selection.

The survey data was analyzed by collecting all responses into a spreadsheet to identify trends and outliers in AI familiarity among respondents. This method is common in mixed-methods research to summarize responses and provide a baseline for further analysis (Kuenzi, 2025). In this study, measures such as frequency, mean value, and standard deviation were used to describe the distribution of responses in categories such as perceived AI usefulness.

The qualitative data from open-ended survey questions was analyzed using thematic analysis as suggested by Braun and Clarke (2013). They describe thematic analysis as: "a method for identifying themes and patterns of meaning across a dataset in relation to a research question" (Braun & Clarke, 2013, 174). This method was suitable because the responses were short, varied, and descriptive. The analysis process also included generating codes for similar responses and compiling them into themes that are relevant to the research question. The themes found from the analysis are listed in Table 4 below.

Table 4: Themes from thematic analysis of the survey

Theme	**Description**
Technical limitations	Respondents found that AI tools do not produce complete answers without context.
User readiness and prompting skills	Respondents underscore the importance of prompt, specific training.
Ethical concerns	Respondents highlight the need for human oversight.

Based on the survey analysis, user readiness and prompting skills emerged as the most prominent themes for testing. For that, a Master Service Agreement (MSA) was selected. The contract was chosen because it contained multiple revisions, allowing sufficient clause comparison and analysis, and because the contract type is quite common, making it generalizable. The versions of the tested MSA were a template from the purchaser, representing the organization's standard MSA terms, and the final negotiated version, which included all agreed-upon modifications. All confidential business details and company identifiers have been made anonymous for the purpose of this thesis to comply with ethics and data privacy. The tests were also conducted in a local testing environment with no external data transmission.

In the final phase of the process, the test results were analyzed. This analysis included a comparison of responses generated by different models and an evaluation of how prompt types affected the quality and usefulness of the outputs. Two AI models were applied with the same contract examples and different types of prompts. With this, the model's performance was tested under varying conditions. All outputs were documented in a data sheet for direct comparison. In addition to the prompts and response, the datasheet included the model name, prompt type, response time, and the number of follow-up prompts required for a subjectively complete answer.

4.2 Reliability and Validity

The question of general research reliability is reviewed in light of the study's validity and reliability. The validity of the research is based on whether the study has researched what it was set out to research. Validity is often divided into outer and inner validity. Outer validity refers to the extent to which the research can be generalized. The inner validity can be further divided into segments, including content validity, construct validity, and criterion validity. When measuring content validity, the question is whether the indicators and concepts used in the study are in line with the theory and whether the concepts cover the phenomenon in question sufficiently. All in all, the validity of research therefore requires reducing systematic errors. This can be achieved by ensuring that participants clearly understand the tasks, such as accurately interpreting questionnaires. (Metsämuuronen, 2006. 55, 64-65)

The reliability of a study can be measured in three ways: repeated measurement, which involves using the same indicators multiple times; parallel measurements, which include using multiple indicators simultaneously; or by measuring internal consistency. The reliability of a study is therefore rooted in the testing-retesting-reliability. This means that reliable research produces similar results when repeated measurements are taken with the same participants, regardless of the researcher conducting the study. For research to be effective, it must be replicable and supported by objective data. (Metsämuuronen, 2006, 65-66)

In the case analysis, reliability was ensured through the use of standardized procedures in testing AI models. The same contract texts and prompts were used across all tests to allow for consistent comparison between prompt types and AI models. This systematic approach makes it possible for another researcher to replicate the tests and obtain results that are generally similar. That being said, it must be acknowledged that AI models are continuously evolving, and their outputs may vary depending on the updated features and capabilities of these models, which poses a challenge for absolute repeatability.

To ensure validity, the survey was conducted to get an understanding of the familiarity, perceptions, and potential use cases of AI among procurement professionals. This provides information on the company's practical readiness to adopt AI into its procurement processes. The case analysis directly tested AI tools on relevant procurement contract tasks, such as clause comparison, contract summarization, and detecting inconsistencies between contract versions. This ensured that the research aligns with the research question of how AI-based language models can be used in supply management.

Inner validity of the thesis was ensured by clearly defining the scope of testing and maintaining consistent conditions across prompt types and models. The prompts were designed to reflect real procurement tasks, thereby strengthening the construct validity. Outer validity is, however, limited due to the fact that the sample size is relatively small and context specific.

AI tools, such as ChatGPT and Grammarly, were utilized throughout the thesis to enhance the grammar, clarity, and coherence of the text, as well as to suggest alternative formulations to avoid repetition. AI was also used as a tool for summarizing survey data and refining the structure of sections to ensure consistency. AI functioned only as a supportive aid to enhance efficiency and readability.

5 Case Analysis

This chapter presents the empirical part of the research. In this chapter, the survey and practical testing are combined to evaluate the use of AI-based language models in procurement. The case analysis follows a logical progression. It begins with interpreting survey results, then with case selection and AI testing. The chapter culminates in an in-depth analysis of prompt performance and a comparison of AI models.

5.1 Survey analysis

To determine how AI-based language models can improve the efficiency of procurement contract work, a survey was conducted among sourcing managers across multiple sourcing departments. The survey aimed to gather information on their experience, current use of AI tools, perceptions of the value AI can bring to their work, and potential barriers to using AI in their daily work.

The familiarity of AI-tools was found to be quite low. When asked about whether they had used AI tools to assist in any phase of the contract life cycle, only four responses (36%) said that they had. The survey also revealed that awareness and training in AI were limited, with six respondents indicating they were unfamiliar and had received no training, three being familiar but untrained, and only two having received any training at all.

In one question, respondents were asked to rate their confidence in evaluating AI-generated contract content for accuracy and legal risks on a scale from 1 to 5. Nine respondents chose the midpoint value of 3, indicating moderate confidence. This aligns with the literature, which emphasizes the necessity of human oversight when integrating AI tools into legally sensitive domains such as contract management (Surden, 2019). In another question, 10 out of 11 participants agreed or strongly agreed that AI use in procurement contracts requires strong human oversight.

Even though many of the respondents had not used AI tools, perceptions were generally positive. The survey showed that 10 (91%) respondents agreed that AI can improve contract drafting efficiency, 9 (82%) agreed that AI can help reduce human errors, and 8 (72%)

believed that AI can assist in identifying compliance and legal risks. The survey also showed that managers with many years of experience rely more on their judgment in making final decisions, with only one respondent fully comfortable relying on AI for contract analysis. This indicates that trust in AI output is still developing and would benefit from transparency and understanding of AI tools.

While there is general openness to AI, the open-ended questions also revealed concerns that could hinder the use and effectiveness of AI tools. These concerns are linked to the themes of technical limitations, user readiness, legal and ethical concerns, and contextual fit with existing practices in procurement.

Respondents questioned the technical limitations in terms of reliability and utility. For example, one respondent noted that "The experiences so far are not promising. AI can only provide generic responses, not specific enough for our contracts". This comment shows that, without the necessary context, AI cannot assist in drafting contracts. Another barrier involves the human side of interactions, as some respondents stated they are unsure how to write effective prompts, which directly affects the quality of AI responses.

"Users seem like the biggest concern. If prompting is done incorrectly and the basic principles of how LLM works are not understood, the results can be bad, and the user may not even realize what is happening."

The statement above, written by a respondent who has a better-than-average understanding of AI tools, points out a common challenge in working with AI tools. While the tools are flexible, they are very sensitive to prompt quality, which also explains, to some degree, why many users have different experiences with using AI.

Multiple respondents also showed concerns with data privacy, copyright, and legal accountability. In heavily regulated areas such as procurement, misinterpretation of contractual clauses or misleading summaries can lead to legal consequences. One respondent noted the question of how reliably AI can translate legal text while preserving obligations.

Together, the concerns from open questions show that improving efficiency in procurement through AI is not only a technical challenge, but a socio-technical transformation. This aligns with the research on successful AI integration and how it requires alignment and reconfiguration of technology, organizational structures, and human practices. (Zhang & Dong, 2024)

All respondents identified specific training areas that would support the adoption of AI tools. The most frequent recommendations were prompt writing and use case-specific training. Several respondents also noted the need to understand ethical boundaries and training on the basic principles of how AI works. The specific use cases mentioned in the survey were: clause drafting and simplification, version comparison of documents, and contract analysis using simpler language. The use case tasks reflect the contract life cycle stages that involve repetitive, time-consuming reading or writing. This makes them ideal for AI.

5.2 Case selection

This chapter explains the reasoning behind selecting contract comparison as the case to be tested in this thesis, specifically through the use of prompting. The main research question of this thesis is how AI-based language models can be used in supply management. With a special focus on improving the efficiency of work related to procurement contracts. Even though AI in procurement can be applied to many cases, as presented in the questionnaire, each case ultimately depends on the quality of the user's input.

The case of contract comparison has been selected as the practical scenario through which different prompting approaches are tested. Contract comparison requires identifying differences in clauses, terms, or legal conditions across two contract versions, making it highly sensitive to the clarity and specificity of the prompt. Using prompting for the test case also directly supports the sub-question of what should be considered when writing a prompt. For example, in contract comparison, a good prompt must specify which clauses or areas should be compared. In drafting, the prompt must specify the language and scope of the document. Therefore, prompting is the foundation for other use cases and can also be transferred across AI tools.

5.3 Testing

The testing of prompts has two main goals. The first goal is to assess how the quality and structure of prompts influence the performance of an AI language model. This includes examining how variations in prompt specificity, clarity, and structure affect the relevance, accuracy, and usefulness of the model's responses. Prompting is the connection between humans and the AI machine, which makes understanding their relationship and how they should interact is essential. As in human communication, clarity, context, and structure determine the quality of understanding.

The second goal of the testing is to identify best practices for writing prompts that increase the usefulness and reliability of AI-generated answers in procurement contract analysis. To achieve this the testing procedure was designed to evaluate the impact of different prompt types on model performance. The prompt styles in Table 5 below ranged from open-ended and straightforward to highly detailed and task-specific.

Table 5: Prompt Variations

Prompt Type	**Description**
Basic Prompt	General instruction without structure (e.g. "Compare contract A and contract B")
Specific prompt	Clearly scoped with clause focus (e.g. confidentiality, liability, termination).
Structured prompt	Use bullet points as steps to guide the AI
Contextual prompt	Include background information and role framing
Layered prompt	Request processes with multiple steps

Each prompt type represents a different level of human-AI interaction. This allows testing to provide insight into how increasing structure and context information affects the analytical accuracy and depth of responses. The testing procedure systematically evaluates the impact of prompt variations on the performance of an AI language model in contract-related tasks.

To assess the outcome of these tests, five evaluation criteria were used. These criteria are shown in Table 6 below. The methodology combines standardized input-output processes, human evaluation, and expert validation to ensure the credibility of the results.

Table 6. Evaluation criteria

Criteria	**Description**
Comprehension	Does the output address the task described in the prompt?
Accuracy	Are the details correctly interpreted?
Detail and Depth	Is the output understandable?
Efficiency	Does the AI highlight meaningful areas?
Handling complexity	How well does AI handle complex prompts?

To ensure that the test is done on a consistent basis, the data had to be consistent. For this case, a procurement contract with multiple critical modifications from the template to the final agreement was selected. Another criterion for the selection was that the type of agreement was commonly encountered by sourcing managers. According to Lee and Palmer (2025), using realistic input materials is critical when evaluating prompt engineering approaches, particularly in settings where AI outputs augment human decision-making. For each prompt variation, the AI language model was instructed to complete predefined contract management tasks. These included summarizing key clauses, identifying potential risks or anomalies, and suggesting edits for clarity. The outputs were systematically collected and documented for analysis.

To further ensure the validity and accuracy of the testing results, the sourcing manager responsible for the contract also evaluated the AI's responses. The manager independently assessed the outputs without knowing which prompt variant had been used to reduce potential bias. The evaluation by the manager focused on the relevance, completeness, accuracy, and clarity of each answer, enabling a systematic comparison.

5.4 Findings from different prompt types

This chapter will go over the empirical findings from the testing of each different prompt type. By comparing the outputs generated by each prompt variation, this chapter aims to reveal how the design of a prompt shapes the behavior of an AI model, thus how prompt engineering can enhance or limit the practical value of AI in procurement work. The findings also provide insights into best practices for human–AI interaction, showing that the effectiveness of contract analysis depends not only on the model's capabilities but also on how effectively the user formulates and structures their prompts. As established in the methodology, five prompt categories: basic, specific, structured, contextual, and layered were tested systematically using identical contract material to ensure comparability. This chapter does not include a comparison of AI models and therefore does not differentiate the models used. The responses to prompts show the general capabilities of AI.

5.4.1 Basic prompt

The basic prompt: "What is the contract about and what are the key differences between contract A and contract B?" required several iterations to yield substantive results. Once provided with the clarification that the documents were a template and a final version, the AI was able to identify the core nature of the agreement and extract the most significant negotiated commercial and legal changes.

Using the basic prompt, the AI was able to correctly identify the documents as a Master Supply Agreement (MSA). It also explained further that it established the legal framework for the Supplier to provide engineering services to the Purchaser via subsequent Work Package Agreements (WPAs) and orders. The AI found a change in the contract's duration. The template specified a fixed term, offering no long-term stability. Whereas the final agreement established an initial term ending at the end of the year, with automatic one-year renewals. This model provides a more stable, ongoing relationship favorable to the Supplier.

The AI also identified a significant reduction in the Supplier's warranty obligation. The template has a 24-month warranty period, which the Supplier negotiated down to 15 months after acceptance or 6 months after delivery to the Purchaser's end-customer. This reduces the Supplier's long-term liability for defects.

The penalty structure for liquidated damages was also modified in the final version of the agreement. The template had a harsh penalty of at least €1,000 per day with no upper limit. In the final version, this was replaced with a more predictable, capped formula: 5% of the relevant Order price per week of delay, with a total cap of 25% of the Order's value. This change substantially reduces the cost for the supplier in the event of delays.

Another crucial finding was with the limitation of liability. The AI identified that the template by the Purchaser contained no financial cap on liability. This clause exposed the Supplier to potentially unlimited damages. The final contract introduced a multi-tiered cap: A General Cap for direct damages at the higher of 100% of the previous year's fees or €150,000. A "Super Cap" for major breaches, such as data protection, at the higher of 200% of the previous year's fees or €3,000,000. This represents a massive concession by the Purchaser and a win for the Supplier. The insurance requirements were also lowered from €5,000,000 per occurrence in the template to just €500,000 in the final agreement.

The changes in Intellectual Property (IP) infringements were not as substantial, but the AI detected a change in IP protection. The template had a broad clause making the Supplier responsible for any third-party IP infringement. This clause was made narrower with additions to the final agreement stating that the Supplier is not liable if the infringement results from the Purchaser’s own modifications to the deliverable or its combination with products not offered by the Supplier.

5.4.2 Specific prompt

The aim of the specific prompt was to give the AI a more specific task. This is a more common way to ask AI for help when the general content of an agreement is known. In this prompt, the AI was asked to: "Compare the delay charge clauses in the contract versions, specifically addressing their scope, duration, and any exceptions." This more focused prompt required the AI to perform a "deep dive" into a single clause.

The AI confirmed that the trigger for penalties in both documents is a "Non-Excusable Delay." However, it precisely identified the change in calculation: the template's flat daily fee of €1,000 was replaced by a percentage-based model of 5% of the Order price per week in the final version. This change made the penalty proportional to the value of the delayed work.

In the template the penalties were allowed to accumulate indefinitely whereas in the final agreement the penalties for delays can not exceed 25% of the price of the impacted order. Both versions of the agreement kept the clause for “Excusable Delays” (e.g., an unavoidable, external event). This is therefore the only way the supplier can avoid penalties in case of a delay.

5.4.3 Structured prompt

The “Structured Prompt” is a multi-step prompt asking to summarize differences in payment terms, identify discrepancies in indemnification, highlight variations in termination, and assess dispute resolution. This prompt required the AI through a logical workflow, giving the user organized and detailed findings.

The key contractual findings with the prompt showed no changes in the payment terms. The AI correctly identified that "60EOM10" payment schedule was not changed. However, it found two subtle, Supplier-favorable, terms that were altered. The first clause that prevents the Supplier from charging for work similar to past projects was removed. The other clause extended the repayment period for unused credit notes from 30 days to 6 months.

The AI found that termination clauses were modified for both parties. On one hand, the Purchaser gained more specific triggers to terminate for delay (e.g., at least five separate Non-Excusable Delay events occur per calendar year). In the other, the Supplier's position was strengthened because its right to terminate was expanded beyond simple non-payment to include any "other material breach" by the Purchaser.

The structured prompt proved itself useful, especially in dispute resolution. Through this prompt, the AI made a critical finding that was missed by more basic prompts. The AI discovered a change in the "performance during a dispute" clause. The template required the Supplier to continue all work pending resolution. The final agreement allows the Supplier to suspend work on the specific "Order under Dispute." This is a massive shift in leverage, as it allows the Supplier to halt a specific project to apply pressure on the Purchaser during a disagreement.

5.4.4 Contextual prompt

The contextual prompt required the AI to take on a requested persona: "You are a lawyer advising a sourcing manager who is considering entering into a agreement. Your task is to compare the original template and the final contract. The manager is particularly concerned about potential liabilities and intellectual property ownership. Please provide a comparison that addresses these areas, identifying which contract version presents a lower risk profile and better protects the IP of the Purchaser."

This prompt required the AI not just to extract data but to compile it into a strategic risk assessment tailored to a specific stakeholder. In this case, the key contractual findings were formatted as a legal memo.

The AI concluded that the original template offered the Purchaser maximum risk protection via uncapped liability. It then translated the contractual changes into a clear business risk: "the new liability caps in the final agreement mean that for most contractual failures, the Purchaser's ability to recover its financial losses is now strictly limited to the defined amounts (€150k or €3m). The reduced insurance requirement was identified as amplifying this risk."

IP Ownership vs. Protection: The AI made a crucial distinction. It confirmed that IP Ownership was secure in both versions: the Purchaser is the sole owner of all newly created "Foreground IP." However, it correctly advised that IP Protection in the form of indemnity was weakened. The new exceptions in the final contract were presented as a significant risk transfer, creating scenarios where the Purchaser could be solely responsible for defending a third-party infringement claim.

5.4.5 Layered prompt

The final layered prompt was a comprehensive, three-step instruction to perform a high-level review, a detailed clause-by-clause comparison, and a final risk assessment (Table 7). This prompt structure yielded the most comprehensive output, effectively combining the findings from all previous prompts into a single, cohesive report.

The results show that the layered prompt enabled the AI model to organize information hierarchically and follow an analytical workflow similar to that of a human analyst. During the first stage, the AI correctly identifies the main contractual elements, including the contractual parties, effective dates, and the overall purpose of the MSA. This established a clear contextual foundation for the following analysis.

In stage 2, the AI model summarized all contractual changes between the template and the final agreement. The analysis included differences in payment terms, liability limitations, and termination clauses, among others, as shown in Table 7. The answers were well structured, factually correct, and demonstrated strong contextual awareness. This stage served as evidence for the risk assessment and recommendations in stage three.

Table 7: Key Contractual Findings (organized by stage):

Stage 1	High level review	Confirmed the parties, the effective date and the overall purpose of the MSA.
Stage 2	Clause-by-Clause Comparison	Summarized all specific changes. Providing a holistic view of the negotiation's outcome. It served as the evidence base for the final stage.
Stage 3	Risk Assessment and Recommendation	The AI combined the clause-level findings into three key strategic risks for the Purchaser in the final agreement. Financial, IP Infringement liability and Operational risk

In the third stage, the AI combines its findings into a strategic summary of the main risk areas affecting the Purchaser. It identified the main risks as financial, intellectual property infringement, and operational. In addition, the model provided a justified rationale for the risks associated with changes to contract clauses.

From an efficiency perspective, the layered prompt required extensive input to craft, but it also significantly reduces the need for follow-up prompts. The AI produces nearly final-quality output with a single prompt. This shows that structured multi-step prompting can significantly save time in contract analysis. In praxis the sourcing manager would also not have to draft the prompt for each contract but could use predefined prompts for each task.

5.5 Comparison of AI models

This chapter compares the responses from both models. The purpose of the comparison is to evaluate how differences in model performance, comprehension, and accuracy affect their suitability for procurement use. The result of this chapter will reveal which model is more suitable for procurement related tasks.

AI model A was consistently accurate across all prompt types. The outputs show that model A correctly identified and compared specific clauses, such as payment terms, liability, and dispute resolution, without misinterpretation. Model B's accuracy is, however, inconsistent. It, for example, provided incorrect information on specific and contextual prompts. A clear example is found in the Specific Prompt. When asked to compare delay charge clauses, Model B's initial response was incorrect. When prompted with a follow-up to correct its answer using the correct document name, it failed, stating: *"I'm sorry, but I'm currently unable to answer questions that require specific document metadata..."*. This shows a critical failure in both accuracy and the ability to self-correct.

Model A's comprehension was at an excellent level. It understood nuance, context, and complex instructions. This was best illustrated in the Contextual Prompt. Model A understood the persona of a "lawyer advising a sourcing manager" and framed its entire response as a legal memo, directly addressing the manager's concerns about liability and intellectual property. For Model B, comprehension was a more difficult task as it struggled with prompts and failed to understand contextual framing. In the same contextual prompt, model B provided a generic comparison without adopting the requested persona or tailoring its analysis to the specific concerns regarding potential liabilities for a sourcing manager. In its response, it simply listed differences rather than giving advice.

In the category of detail and depth, model A was able to provide a detailed, in-depth analysis with implications of the findings." The layered prompt is a prime example. Model A followed the three-stage structure perfectly, not only comparing clauses but also performing a risk assessment and concluding with the remaining risks for the Purchaser. It analyzed the implications of the changes, such as how liability caps create financial risk. Model B’s responses were often high-level and lacked the same kind of depth that model A was able to produce. Model B listed changes, such as the non-compete clause, which was reduced to six months, but generally failed to explain why these matters or what the strategic implications are for the user.

Model A was efficient with more specific and structured prompts." The processing times listed show model A was often faster (45 seconds for the structured prompt vs. model B’s 1 minute 10 seconds). More importantly, its high accuracy and complete answers meant no follow-up prompts were needed for complex queries. Model B required follow-ups on the basic prompt and was unable to self-correct. The need for additional prompts to get a correct or complete answer, combined with its failure on the specific prompt, makes its overall process inefficient.

Model A excelled with structured and layered prompts, accurately following multi-step instructions. The layered prompt required a multi-stage analysis (Initial Review → Clause Comparison → Risk Assessment). Model A’s output followed this structure, which showed its ability to manage a complex workflow within a single prompt.

Unfortunately for model B, it also struggled with multi-step instructions, leading to incomplete or inaccurate answers. The data indicates that as the prompts became more complex by adding instructions such as "what are the differences," model B’s ability to provide a useful response degraded significantly. Model B’s response to the layered prompt was a critical failure, incorrectly concluding that both versions presented a "similar risk profile," advice that would have been highly misleading to a sourcing manager.

5.6 Summary of the testing

The testing aimed to evaluate the performance and capabilities of two AI-based language models and the effectiveness of different prompt types for procurement contract analysis-related tasks. Both models were tested using five prompt styles (Table 8). The prompt styles were: Basic, Specific, Structured, Contextual, and Layered. Testing followed a progressive logic, with each prompt becoming more comprehensive. The basic prompt examined the models' general understanding and factual accuracy, whereas the specific prompt tested how well the Model was able to focus on individual clause details. The structured prompt introduced step-by-step instructions that required the Model to follow directions more strictly. The aim of the contextual prompt was to have AI imitate the professional role of a legal advisor. Finally, the layered prompt required the Models to follow detailed, multi-step processes similar to human processing.

Table 8: Prompt styles and prompts

Prompt style	**Prompt**
Basic	What is the contract about, and what are the key differences between contract A and contract B?
Specific	Compare the delay charge clauses in the contract versions, specifically addressing their scope, duration, and any exceptions.
Structured	Compare the contract template and the final contract by following these steps: Summarize the key differences in their payment terms. Identify any discrepancies in the indemnification clauses. Highlight variations in the termination rights for each party. Assess which contract offers more favorable dispute resolutions for the Purchaser
Contextual	You are a lawyer advising a sourcing manager who is considering entering into an agreement. Your task is to compare the original template and the final contract. The manager is particularly concerned about potential liabilities and intellectual property ownership. Please provide a comparison that addresses these areas, identifying which contract version presents a lower risk profile and better protects the IP of the Purchaser.
Layered	Perform a comprehensive comparison of the contract template and the final contract, following these stages: Stage 1: Initial Review and High-Level Assessment Read both versions thoroughly to understand the overall purpose and scope. Identify the parties involved and the effective dates for each contract. Note any obvious differences in their structure or length. Stage 2: Clause-by-Clause Comparison of Key Provisions Scope of Services/Goods: Compare the detailed descriptions of what is being provided. Payment Terms: Analyze currency, payment schedules, late payment penalties, and invoicing procedures.

	Liability and Indemnification: Detail the limitations of liability, disclaimers, and indemnification obligations. Intellectual Property: Compare clauses related to ownership, licensing, and infringement. Confidentiality: Examine the definition of confidential information, duration of obligations, and permissible disclosures. Termination: Identify conditions for termination, notice periods, and post-termination obligations. Governing Law and Dispute Resolution: Compare the chosen jurisdiction and methods for resolving disputes. Stage 3: Risk Assessment and Recommendation Based on your clause-by-clause comparison, identify the key risks associated with the final version for the Purchaser Formulate a conclusion regarding the remaining risks for the Purchaser in the final version.

The analysis revealed a significant performance gap between Model A and Model B. As seen in Table 9, Model A consistently outperformed Model B across all evaluated criteria, particularly in handling complex, structured, and contextual prompts. Model A demonstrated a superior ability to comprehend nuances, analyze data, and structure its output according to user specifications. This makes it a reliable and effective tool for detailed tasks, such as contract clause comparison, risk identification, and summarizing legal documents. In contrast, Model B produced more generic and inconsistent replies. It performed acceptably on basic tasks, but it struggled with accuracy, depth, and following multi-step instructions. This resulted in inconsistent responses that required user intervention, such as follow-up prompts or corrections, to reach an acceptable level of quality

A key differentiation factor between the two models was their handling of structured and layered prompts. Model A's performance improved when the level of structure improved, and it benefited from task breakdowns, role framing, and contextual information. This aligns with existing research by Liu et al. (2023), who explain that advanced LLM's leverage structured input to produce more accurate responses. Model B's performance did not improve significantly when provided with a detailed structure. It often ignores individual subtasks or produces incomplete responses. This showed its limitations in reasoning capabilities and ability to understand instructions. Overall, the results show that Model A is significantly better suited for procurement contract management, because Model B's limitations restrict it to basic tasks.

Table 9: Comparison of AI models

Feature	AI model A	AI model B
Comprehension	Great understanding of context and complex instructions.	Struggled with complex prompts and failed to grasp contextual framing.
Accuracy	Consistently high across all prompt types.	Inconsistent; provided incorrect information on specific and contextual prompts.
Detail & Depth	Provided detailed, in-depth analysis with implications of the findings.	Responses were often high-level and lacked analytical depth.
Efficiency	Required multiple prompts for the basic query but was highly efficient with more specific and structured prompts.	Required follow-up on the basic prompt and was unable to self-correct when its initial answer was wrong.
Handling Complexity	Excelled with structured and layered prompts, following multi-step instructions accurately.	Struggled with multi-step instructions, leading to incomplete or inaccurate answers.

6 Discussion and conclusion

This chapter discusses the main findings in relation to the theoretical framework and research questions. The thesis examined how AI-based language models can support supply management, particularly in procurement contract management. The thesis was done using a mixed-methods approach combining a survey and prompt-based testing. Each research question and its corresponding answer are discussed separately in this section of the thesis. The chapter begins with a summary of the findings, after which the main research question is answered, followed by the answers to the sub-questions.

The findings in this thesis show that AI-based language models can support supply management functions such as procurement contract management when used as decision-support tools rather than fully automated systems. The survey revealed that sourcing managers experience areas in contract review work that are somewhat inefficient and could be improved through AI. The empirical testing showed that AI models can address these challenges by summarizing clauses, identifying differences between contract versions, and highlighting potential risks. Through the study, prompt engineering and context engineering were important factors that influenced output quality. This confirmed prior research by Rajasekaran et al. (2025) that well-designed prompts act as the mechanism through which human expertise outperformed Model B in comprehension, accuracy, depth, efficiency, and handling of complex instructions, making it more suitable for contract analysis tasks. Overall, the results show that AI-based language models can be effectively applied in procurement contract management if human expertise, prompt design, and contextual understanding are integrated. This supports the view that procurement activities are knowledge-intensive and can benefit from automation (Van Weele, 2018).

6.1 Discussion on findings

The main research question addresses the overarching purposes of this thesis: How can AI-based language models be used in Supply Management? This thesis concludes that AI-based language models can be used in supply management as decision-support tools. AI models add value by combining automation, contextual understanding, and human expertise. The most significant value lies in improving efficiency, risk visibility, and communication with internal stakeholders such as project managers and legal. However, as Zhang and Dong (2024) emphasize, the effectiveness of AI in supply management depends on human oversight and alignment with organizational context, confirming that AI's role is complementary rather than substitutive.

To understand how AI-based language models can be used in supply management, this thesis combines procurement and contract management frameworks (APM, 2017; Garrett & Rendon, 2012) with natural language processing, prompt engineering, and human-computer interaction. These theories explain how AI can transform tasks such as contract analysis from manual to automated processes. The practical testing phase demonstrated that language models can analyze complex legal documents, summarize differences, and identify risks when guided by structured, layered prompts. Therefore, AI-based language models also contribute to the digital transformation of supply management by improving data accessibility and standardizing analysis.

Even though the case analysis shows that AI can help in supply management processes significantly, it still relies heavily on human interaction through prompt design and oversight. As seen in the study, poorly structured prompts can lead to incomplete and or even inaccurate results. Therefore, it is important to provide users with training in creating structured, contextual, and layered prompts. This is also in line with Liu et. al.'s (2023) research on prompt engineering. Furthermore, AI cannot be used as a substitute for professionals, especially in legally sensitive contexts, as professionals still retain final accountability for contracts.

Sub question 1: How can AI-based language models be used to improve the efficiency of work related to procurement contracts? This sub-question focuses on one of the main issues highlighted by sourcing managers in the survey. The testing demonstrated clear differences in how AI performs contract-related tasks depending on prompt quality. The following paragraphs discuss how these findings translate into concrete efficiency improvements in contract-related work.

The findings in regard to sub-question 1 demonstrate that AI-based language models can enhance the efficiency of procurement contract management. AI models can do that by automating repetitive tasks, enhancing accuracy in contract comparison, and supporting in risk assessment. The efficiency gains found in this thesis are in line with the broader literature on AI in procurement processes, which highlights the potential of AI to reduce manual effort (Guida et. al. 2023; Mittal et al. 2024).

The automation of manual tasks provides sourcing managers with notable efficiency gains. Activities such as clause comparison, summarizing contracts, and identifying modifications require long human review, but with a few prompts, they can be done in a matter of minutes. Dilmegani (2025) stated in his article about use-cases for AI in Procurement that AI can streamline procurement processes. This became clear in the case analysis as the layered prompts enabled AI to create a report consisting of a summary, clause-by-clause comparison, and a risk assessment.

Manual contract review is prone to human errors due to the complexity and length of contracts. That is where AI can help improve accuracy and reduce human error. The tests confirmed that AI was able to identify subtle yet impactful changes in the contract versions that could be overlooked during manual review. This finding supports Surden (2019), who emphasizes the role of AI in improving consistency and reliability in legal text analysis. The contextual prompts in the case analysis show that AI can translate contract changes into business risks, thereby supporting decision-making. Zhang and Dong (2024) argue that integrating AI is most effective when it complements human judgment. In this case, it does that by turning clauses into strategic advice, accelerating decision-making. To conclude the answer to the first sub-question. AI-based language models are a valuable tool for improving efficiency in procurement by significantly reducing the workload of review tasks, minimizing the likelihood of human errors, and generating decision-ready insights.

The second sub-question is ”How can AI models help identify risks, inconsistencies, or non-compliance in procurement contracts?” The study showed that AI models can provide important support in identifying risk, inconsistencies, and potential non-compliance within procurement contracts, supporting Surden’s (2019) argument that AI can improve legal consistency and accuracy. The AI models tested in the case were able to detect differences between templates and final agreements and assess their implications for sourcing managers. Throughout the testing of contextual and layered prompts, the AI was able to translate the findings into concrete risks and explain how those change the Purchaser's ability to, for example, recover losses. This also aligns with Rendon and Garrett (2016), who highlight that many contract disputes stem from undocumented or unclear changes.

While AI successfully identified risks and inconsistencies, its accuracy varied depending on the model used and the quality of the prompt. One model that was tested misinterpreted clauses, which could have led to misleading conclusions if the user had no prior knowledge of the contract. To conclude the second sub-question, AI models are a great aid in reducing the manual effort required for contract analysis. Its usefulness depends heavily on prompts and human validation of answers.

The third sub-question: ”What to take into consideration when writing a prompt? Including prompt engineering and context engineering” stems from the importance of Human-AI interaction. While AI models have strong processing capabilities, their effectiveness depends largely on how users communicate with them (Dix et al. 2004). The findings of the case analysis show that the effectiveness of AI models in procurement contract analysis is highly dependent on the way prompts are made. Prompts act as the interface between the user and the model, and their quality directly affects the relevance, accuracy, and usefulness of answers (Knoth et. al 2024). Key considerations when writing a prompt include the structure, clarity, and depth of the prompt.

The study also found that basic prompts usually only produce surface-level answers and require additional prompts to get useful information. In addition, a basic prompt may miss critical information because the model does not know the specifics of the users request. Without specificity, the efficiency of using AI for procurement is reduced. To further improve the quality of the answer, the user should add context and a role definition to the prompt. By adding context, the responses changed from generic to specific, with strategic implications. This addition to the prompt suggests that prompts should not only be directed at a certain task but also clarify the intended use and perspective of the answer. This confirms theories of prompt and context engineering (Liu et al., 2023; Koch et al., 2025), suggesting that prompt design should be viewed as a communicative and cognitive skill that translates human intent into structured instructions for AI.

When prompting larger tasks, it was found useful to add structured guidance to the prompt. By dividing the request into steps, the AI was able to follow a logical workflow, which made the answer easier to comprehend. For the most comprehensive results, the user must also include depth and layering. With this, the AI was able to produce a final answer that was like what a human analyst would be able to produce. By identifying risks that are not evident in the contract, the AI can add real value to the user and not just save time.

6.2 Theoretical and practical implications

The theoretical findings of this thesis suggest that the use of AI tools in supply management and procurement contract management represents an intersection between traditional working methods and human-AI interaction. The thesis also contributes to the theoretical understanding of how artificial intelligence can improve efficiency in contract management. It extends the existing framework for contract management by APM (2017) and Garret and Rendon (2012) by integrating prompt engineering as a new dimension for supply management. The results of the thesis also show that the most valuable efficiency improvements come from the interaction between human expertise and AI capabilities. This supports the social perspective of Zhang and Dong (2024) and aligns with the findings of Balkan and Akyüz (2025), whose study showed that decision-making in supply management holds significant potential for AI applications.

The testing also contributed to theoretical insights into how prompting acts as the mechanism between human expertise and AI reasoning. This supports the findings of Gutheil et al. (2025), who demonstrated through a study of test groups that users with guidance on how to use AI consistently outperformed the control group which did not get guidance. In addition to human-AI interaction, this study also contributes to the understanding of context engineering in prompting. The testing showed that the response quality improved when more context was introduced. This aligns with Koch et al. (2025), who argue that the quality of responses increases when humans design the context in which AI reasoning occurs.

The practical implications of the thesis show that AI can be an efficient tool for professionals in supply management, especially in procurement contract management. In practice, AI models can assist in drafting, reviewing, and comparing contracts, which reduces the manual workload. To achieve these efficiency gains, the findings emphasize the importance of context and prompt quality through user training. For the subject company, this means the most effective way to increase valuable AI use is to invest in prompt and context engineering training and in teaching best practices for AI use in procurement. Another way to improve how managers use AI is to provide standardized prompt templates. The risk with templates is, however, that they can lack specificity, which is an important part of effective prompts.

For further AI development, this thesis shows that current models perform well in specific tasks such as summarizing and comparing. As stated in this thesis, the performance is heavily dependent on the quality of prompts and contextual information. This means that future AI development could focus on designing models that are able to fill the gaps in prompts without extensive user input. Models could also be developed more towards being an assistant by having it ask questions about missing context or ask reassuring questions about what has been prompted.

6.3 Limitations and future research

The findings of this thesis provide insight into the use of AI-based language models in procurement contract management. The scope of supply management, however, is so broad that several limitations must be acknowledged. The scope of the empirical testing was limited to a small number of prompts and two AI models. While this enabled a detailed analysis, it means the results cannot be generalized to all language models or contexts, as different AI systems may vary significantly in their capabilities.

Another limitation is that the testing case data was drawn from a single contract type. This ensured that the data was consistent but also narrowed the practical scope. In addition, the thesis focused primarily on prompt engineering and textual output quality, therefore leaving out factors such as data privacy, security, and ethical considerations. These aspects were discussed in the theoretical framework of the thesis to acknowledge their importance for responsible AI use but were intentionally left out in the empirical testing.

Future research should build upon these findings by expanding the empirical and theoretical limitations of AI in supply management. Future studies could use a larger and more diverse set of contracts across multiple industries and legal systems to better understand how AI models handle varying factors. Future studies could also test different types of AI systems. As AI’s usage increases in organizations, future research can utilize more data on long-term effects such as cost reduction and compliance improvements.

References

Amershi, S., Weld, D., Vorvoreanu, M., Fourney, A., Nushi, B., Collisson, P.,Horvitz, E. (2019) Guidelines for human–AI interaction. Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, 1–13. Available at: https://doi.org/10.1145/3290605.3300233 Accessed: 13.10.2025

APM (2017) APM Guide to Contracts and Procurement. 1st edition. Princes Risborough: Association for Project Management.

Balkan, D., Akyüz, G.A. (2025). Artificial intelligence (AI) and machine learning (ML) in procurement and purchasing decision-support (DS): a taxonomic literature review and research opportunities. Artif Intell Rev 58, 341

Braun, V. & Clarke, V. (2013) Successful qualitative research : a practical guide for beginners. Los Angeles, CA: SAGE.

Celik, I. (2023). Towards Intelligent-TPACK: An empirical study on teachers' professional knowledge to ethically integrate artificial intelligence (AI)-based tools into education. Computers in Human Behavior, 138, 107468.

Chen, L. et al. (2022) Artificial intelligence adoption in business-to-business marketing: toward a conceptual framework. The Journal of business & industrial marketing. [Online] 37 (5), 1025–1044.

Creswell, J. W. & Plano Clark, V. L. (2018) Designing and conducting mixed methods research. Third Edition. Thousand Oaks, Calif: Sage Publications.

Dilmegani, C. (2025) 10 AI Procurement Use Cases & Case studies in 2025. AI Multiple Research. Available at: https://research.aimultiple.com/ai-procurement/ Accessed: 4.6.2025

Dix, A., Finlay, J., Abowd, G., & Beale, R. (2004). Human–computer interaction (3rd ed.). Pearson.

Edkins, A., Geraldi, J., Morris, P., & Smith, A. (2013). Exploring the front-end of project management. Engineering Project Organization Journal, 3(2), 71–85. Available at: https://doi.org/10.1080/21573727.2013.775942 Accessed: 13.10.2025

Gao, Y. et al. 2023. Retrieval-Augmented Generation for Large Language Models: A Survey. Shanghai Research Institute for Intelligent Autonomous Systems

Gong, C. & Ribiere, V. (2021) Developing a unified definition of digital transformation. Technovation. [Online] 102102217-.

Garrett, G. & Rendon, R. (2012). The Contract Changes Management Process: Managing and Controlling Contract Changes. Contract Management. 52. 56-64.

Gerber, J. (2025). Top 10 AI Procurment Tools in 2025 – Use Cases, Examples & Key Results. Suplari [Online] Available at: https://suplari.com/blog/top-10-ai-procurement-tools/ Accessed: 16.10.2025

Guida, M. et al. (2023) The role of artificial intelligence in the procurement process: State of the art and research agenda. Journal of purchasing and supply management. [Online] 29 (2), .

Guidotti, R. et al. (2018) A Survey of Methods for Explaining Black Box Models. ACM computing surveys. [Online] 51 (5), .

Harju, A., Schaëfer, K., Hallikas, J., Kähkönen, A-K. (2024) The role of risk management practices in IT service procurement: A case study from the financial services industry, Journal of Purchasing and Supply Management, Volume 30, Issue 2, 100899, ISSN 1478-4092, Available at: https://doi.org/10.1016/j.pursup.2024.100899. Accessed: 13.10.2025

Fenwick, A. et al. (2024) The critical role of HRM in AI-driven digital transformation: a paradigm shift to enable firms to move from AI implementation to human-centric adoption. Discover Artificial Intelligence. [Online] 4 (1), 34–16.

Ha, S. et al. (2024) Guided By AI: Navigating Trust, Bias, and Data Exploration in AI-Guided Visual Analytics. Computer graphics forum. [Online] 43 (3), .

Holdsworth, J. (2023). What Is AI Bias? [online] IBM. Available at: https://www.ibm.com/topics/ai-bias.

Idelson, P. (2025) State of AI in Procurement in 2025. Art of Procurement. Availabe at: https://artofprocurement.com/blog/state-of-ai-in-procurement#:~:text=Image%3A%20percent%20of%20procurement%20professionals,using%20generative%20ai Accessed: 2.6.2025

Irshadullah, A.M., (2024) Supplier Relationship Management: Strategies for Enhancing Collaboration and Performance. Scribe and scroll Publishing ltd Dimension of Discovery: THE MULTIDISCIPLINARY RESEARCH PARADIGM (pp.279)

Joyce, W. B. (2006) Accounting, purchasing and supply chain management. Supply chain management. [Online] 11 (3), 202–207.

Kamyshanskaya, S. (2021) AI and Data Privacy: Managing Risk in the Boardroom. Board leadership. [Online] 2021 (174), 6–7.

Kang, Y. et al. (2020) Natural language processing (NLP) in management research: A literature review. Journal of management analytics. [Online] 7 (2), 139–172.

Kananen, P., & Puolitaival, H. (2019) Tekoäly bisneksen uudet työkalut. Helsinki. Alma Talent.

Kolosova, T. & Berestizhevsky, S. (2021) Supervised machine learning : optimization framework and applications with SAS and R. 1st ed. Boca Raton, Florida ; CRC Press. 9-10

Kim, J.-I. & Shunk, D. L. (2004) Matching indirect procurement process with different B2B e-procurement systems. Computers in industry. [Online] 53 (2), 153–164.

Klijn, E. H., Metselaar, S., & Warsen, R. (2023). The effect of contract- and network management on performance and innovation in infrastructure projects. Public Money & Management, 44(5), 428–437. Available at: https://doi.org/10.1080/09540962.2023.2204533 Accessed: 13.10.2025

Knoth, N., Tolzin, A., Janson, A., Leimeister, J.M. (2024) AI literacy and its implications for prompt engineering strategies, Computers and Education: Artificial Intelligence, Volume 6, 100225

Koch, D. et al. (2025) Prompt Engineering in the Enterprise – An Introduction : Competitive Advantages through Generative AI and Large Language Models. 1st ed. 2025. [Online]. Wiesbaden: Springer Fachmedien Wiesbaden.

Korhonen, V. (2009) Tapauksena tapaustutkimus. Aikuiskasvatus. [Online] 29 (1), 66–67.

Korzyński, P. et al. (2023) Artificial intelligence prompt engineering as a new digital competence: Analysis of generative AI technologies such as ChatGPT. Entrepreneurial Business and Economics Review. [Online] 11 (3), 25–38.

Kuenzi, M. (2025) How to analyze data from a survey and unleash the full potential of your research. Surveycto. [Online] Available at: https://www.surveycto.com/analysis-reporting/analyze-data-from-survey/ Accessed: 17.10.2025

Kristensen, T. S. (2021) ‘AI in Business and Education’, in Artificial Intelligence: Models, Algorithms and Applications. [Online]. United Arab Emirates: Bentham Science Publishers. p. 71-72

Lai, J. et al. (2023) Large Language Models in Law: A Survey. arXiv.org.

Lambert, D. & Schwieterman, M. (2012) Supplier relationship management as a macro business process. Supply Chain Management: An International Journal 17, 3, 337- 352.

LeCun, Y. et al. (2015) Deep learning. Nature (London). [Online] 521 (7553), 436–444.

Lee, D. & Palmer, E. (2025) Prompt engineering in higher education: a systematic review to help inform curricula. International Journal of Educational Technology in Higher Education. [Online] 22 (1), 7–22.

Litera n.d. Contract reviews powered by AI. Litera. Available at: https://www.litera.com/products/kira Accessed: 10.4.2025

Liu, P. et al. (2023) Pre-train, Prompt, and Predict: A Systematic Survey of Prompting Methods in Natural Language Processing. ACM computing surveys. [Online] 55 (9), 1–35.

Monczka, R. M. et al. (2014) Purchasing and supply chain management. 6th ed. Boston, MA: Cengage Learning.

Metsälehto, J. (2006). Tutkimuksen tekemisen perusteet ihmistieteissä. 4. painos Vajaakoski. International Methelp KY

Merilehto, A. (2018) Tekoäly: matkaopas johtajalle. Helsinki. Alma Talent

Mwesiumo, D., Nujen. B., Buvik. A. (2021) Driving collaborative supply risk mitigation in buyer-supplier relationships. Supply Chain Forum. 22(3). Pages 1-13.

Nandy, A. et al. (2018) Reinforcement Learning With Open AI, TensorFlow and Keras Using Python. [Online]. Berkeley, CA: Apress.

Norman, D. (2013). The Design of everyday things. Revised and expanded edition. [Online] New York. Basic books.

Osmonbekov, T. et al. (2024) Artificial intelligence advancements in procurement: transforming organizational buying behavior. The Journal of business & industrial marketing. [Online] 39 (12), 2745–2758.

Park, J., Shin, K., Chang, T. W., & Park, J. (2010) An integrative framework for supplier relationship management. Industrial Management & Data Systems.

Rajasekaran, P., Dixon, E., Ryan, C., Hadfield, J. (2025) Effective context engineering for AI agents. Anthropic. [Online] Accessed: 07.10.2025. Available at: https://www.anthropic.com/engineering/effective-context-engineering-for-ai-agents

Reynolds, S. (2024) Exploring the Impact of Collaborative Practices on Supply Chain Resilience. Preprints.org

Santos, R., Silva, M., & Costa, P. (2025). Supplier risk in supply chain risk management: An integrated framework. Applied Sciences, 15(13), 7128.

Sabri, H. et al. (2025) Performance of three artificial intelligence (AI)-based large language models in standardized testing; implications for AI-assisted dental education. Journal of periodontal research. [Online] 60 (2), 121–133.

Schwartz, H. M. (2014). Multi-agent machine learning : a reinforcement approach. 1st edition. Hoboken, New Jersey: John Wiley & Sons, Inc.

Sollish, F. (2011). Strategic global sourcing best practices. Hoboken, NJ: Wiley.

Srivastava S. K. et al. (2024) Exploring the Potential of Large Language Models in Supply Chain Management: A Study Using Big Data. Journal of global information management. [Online] 32 (1), 1–29.

Surden, H. (2019) Artificial Intelligence and Law: An Overview, 35 Ga. St. U. L. Rev. 1305 (2019), available at https://scholar.law.colorado.edu/faculty-articles/1234.

Swink, M. et al. (2020) Managing operations across the supply chain. Fourth edition. New York: McGraw Hill.

Tarigan, Z.J.H., Siagian, H., Basana, R.S., and Jie, F., (2019). Effect of key user empowerment, purchasing strategy, process integration, production system to operational performance, E3S Web of Conferences, 130, 01042, https://doi.org/10.1051/e3sconf/201913001042.

Wasthy, N. et al. (2022) 'The effect of purchasing strategy on supply management performance through supplier integration and supplier relationship management', in Siska Noviaristanti (ed.) Contemporary Research on Business and Management. 1st edition [Online]. United Kingdom: CRC Press. pp. 241–245.

Wieland, D. & Ivens, B. (2025) Not quite alike: Supplier relationship management in B2B marketing and supply chain literature.Industrial Marketing Management. Volume 126. Pages 30-43

Wohlin, C. (2021) Case Study Research in Software Engineering—It is a Case, and it is a Study, but is it a Case Study? Information and software technology. [Online] 133106514-.

Yin, R.K. (2003). Case study research. Design and methods. 3rd edition. London: Sage.

Zhang, Y. & Dong, C. (2024) Exploring the Digital Transformation of Generative AI-Assisted Foreign Language Education: A Socio-Technical Systems Perspective Based on Mixed-Methods. Systems (Basel). [Online] 12 (11), 462-.

Zou, W. et al. (2019) The influences of contract structure, contracting process, and service complexity on supplier performance. International journal of operations & production management. [Online] 39 (4), 525–549.

Appendix 1:

AI for Procurement – Survey

Section 1: Background Information

1. How many years of experience do you have working in procurement?

2. What types of contracts/paperwork do you primarily work with?

3. Which Sourcing Department are you from?
   - Project sourcing
   - Industrial sourcing
   - General procurement
   - Other: _______

Section 2: Current Use of AI

4. Have you used AI tools to assist with managing contracts in any phase of the contract life cycle?
   - Yes
   - No

5. How familiar are you with AI and have you received any training related to using it?
   - Not familiar and no training
   - Not familiar but have received training

- Familiar and no training
- Familiar and have received training

6. Which of the following tasks have you performed using AI? (Select all that apply)
   - Generating summaries of contracts
   - Extracting key terms and obligations from contracts
   - Translating contract clauses into simpler language
   - Checking for inconsistencies in contract wording
   - Drafting emails or reports
   - Other: _______

7. How confident are you in evaluating AI-generated contract content for accuracy and legal risks?

Section 3: Perspectives on AI's Potential

8. To what extent do you agree with the following statements?
   (Scale: 1 – Strongly Disagree, 5 – Strongly Agree)

9. AI language models can significantly improve contract drafting efficiency

10. AI can help reduce human errors in contract management

11. AI tools can help identify risks and compliance issues in contracts

12. The use of AI in procurement contracts requires strong human oversight

13. I would feel comfortable relying on AI for analyzing contracts

14. Please formulate a question about the content of a specific contract and its relation to another contract. Also include your thought process and the important things you focus on when writing a prompt.

Section 4: Future Adaptations

15. What are the biggest challenges or concerns regarding the use of AI in procurement contract management?
 - Training in writing prompts
 - Understanding of how AI tools work
 - Ethics and best practices for responsible AI use
 - Use case-specific training
 - Other: _______

16. Would you be open to using AI-powered tools for daily work if they were proven to be reliable and secure?
 - Yes
 - No
 - Yes, but with caution

17. In which specific use cases, related to managing contracts (e.g., drafting), would you like to see AI implemented?

18. What kind of training or support would help you use AI language models more effectively in your work?

19. Do you have any additional comments on the use of AI in procurement related to managing contracts?