Vaasan yliopisto
UNIVERSITY OF VAASA

Kastari Lassi

Artificial Intelligence in Procurement: Leveraging Google Gemini to Improve the Purchasing Process

School of Technology and Innovations
Master’s thesis in Economics and Business Administration
Master’s Degree Programme in Industrial Management

Vaasa 2025

UNIVERSITY OF VAASA
School of Technology and Innovations

Author:	Kastari Lassi
Title of the Thesis:	Artificial Intelligence in Procurement: Leveraging Google Gemini to Improve the Purchasing Process
Degree:	Master of Science in Economics and Business Administration
Programme:	Master's Programme in Industrial Management
Supervisor:	Jyri Naarmala
Year:	2025 **Number of pages:** 75

ABSTRACT:

This study explores the application of Google Gemini AI within the technical services purchasing process at a case company. The procurement department at the case company has identified a growing need to explore and implement AI-driven solutions to improve the efficiency and quality of its operations. While a growing body of literature demonstrates the potential of AI to enhance various facets of the purchasing process, research in this area is still developing. This research employs a design research approach, integrating Lean Six Sigma methodology, and uses data collected through observations, documentation, and workshops. The analysis includes lead time and defect measurements from five runs for each pilot case, comparing old and new process designs. Findings indicate that Google Gemini can reduce lead times for PO text and RFQ text generation; however, it also introduces challenges, including difficulties in adhering to templates and ensuring output consistency. The study reveals that while AI offers potential for efficiency gains in procurement, current limitations necessitate careful consideration of its implementation and continued development.

KEYWORDS: Procurement; Artificial Intelligence; Purchasing; Lean Six Sigma; Design Research; Process Development; Large Language Models

VAASAN YLIOPISTO
Tekniikan ja innovaatiojohtamisen yksikkö

Tekijä:	Kastari Lassi
Tutkielman nimi:	Artificial Intelligence in Procurement: Leveraging Google Gemini to Improve the Purchasing Process
Tutkinto:	Kauppatieteiden maisteri
Oppiaine:	Tuotantotalouden maisteriohjelma
Työn ohjaaja:	Jyri Naarmala
Valmistumisvuosi:	2025 **Sivumäärä:** 75

TIIVISTELMÄ:

Tämä tutkimus tarkastelee Google Gemini tekoälyn soveltamista teknisten palveluiden ostoprosessissa eräässä case-yrityksessä. Yrityksen hankintaosasto on tunnistanut kasvavan tarpeen tutkia ja ottaa käyttöön tekoäly-pohjaisia ratkaisuja toimintojensa tehokkuuden ja laadun parantamiseksi. Vaikka kasvava määrä kirjallisuutta osoittaa tekoälyn potentiaalin parantaa hankintaprosessin eri osa-alueita, tutkimus tällä alueella on vielä kehittymässä. Tämä tutkimus käyttää design research -lähestymistapaa, jossa yhdistetään Lean Six Sigma -metodologia, ja siinä hyödynnetään havaintojen, dokumentaation ja työpajojen avulla kerättyä dataa. Analyysi sisältää läpimenoajan ja vikojen mittaukset viidestä ajosta kullekin pilottitapaukselle, vertaillen vanhoja ja uusia prosessimalleja. Tulokset osoittavat, että Google Gemini voi lyhentää ostotilaustekstin ja tarjouspyyntötekstin luontiin kuluvaa aikaa; samalla siinä on kuitenkin haasteita, kuten mallipohjien noudattamisessa ja tuotoksen yhdenmukaisuuden varmistamisessa. Tutkimus osoittaa, että vaikka tekoäly tarjoaa potentiaalia tehostaa hankintoja, nykyiset rajoitukset edellyttävät huolellista harkintaa sen käyttöönotossa ja jatkuvaa kehitystä.

AVAINSANAT: Hankinta; Tekoäly; Osto; Lean Six Sigma; Design Research; Prosessikehitys; Suuret kielimallit

Contents

Pictures

Figures

Tables

1 Introduction

The application of Artificial Intelligence (AI) within procurement represents a young yet rapidly expanding field of study. While research in this area is still developing, a growing body of literature demonstrates the potential of AI to enhance various facets of the purchasing process. At the case company, the procurement department has identified a growing need to explore and implement AI-driven solutions to improve both the efficiency and quality of its operations. The case company has recently acquired licenses for the Google Gemini AI tool, necessitating a thorough understanding of how to leverage its capabilities effectively.

This study aims to identify potential process improvements through the Google Gemini, with the expectation that the findings will translate into tangible benefits within the company's existing purchasing processes. Google Gemini is the tool that is used in the study and there are many different purchasing processes at the case company, but this study is limited to the technical services purchasing process for projects. However, there are similarities in different purchasing processes and results of this study could be in certain respects be extended to other purchasing processes as well. For science this study aims to produce new information on practical capabilities of large language models, specifically within the context of purchasing processes utilizing Google Gemini.

The research examines how Google Gemini can enhance the technical services purchasing process at the case company, which serves as the central research question. To address this question, the study sets out several research objectives. First, it aims to identify sources of waste and defects within the current purchasing process and assess how Google Gemini could help reduce these inefficiencies. Next, it seeks to identify potential pilot applications of Google Gemini within the process and select the most promising candidates for deeper analysis. Finally, the study will define, implement, measure, and evaluate the impact of Google Gemini in the selected pilot cases.

2 Literature review

This chapter presents a review of the relevant literature underpinning this research. This review is structured to explore key concepts related to purchasing process, artificial intelligence (AI) and Lean Six Sigma (LSS).

It begins by examining the traditional purchasing process, including specific focus on the nuances of technical services purchasing and process in the industrial management context. The chapter then transitions to an exploration of artificial intelligence, tracing its historical development and delving into the current landscape of Large Language Models (LLMs), with particular attention to Google Gemini. Following this, the review investigates the field of AI in procurement, analyzing its various applications, documented benefits, and potential challenges. Finally, the chapter concludes with a review of Lean Six Sigma, analyzing the core principles of Lean and Six Sigma individually, and examining the advantages and disadvantages associated with their combined implementation.

2.1 Procurement and Purchasing Process

Procurement is a strategic business function that ensures the identification, sourcing, and management of external resources essential for achieving organizational objectives. It explores supply market opportunities and implements resourcing strategies to optimize outcomes for the organization, stakeholders, and customers. It is a proactive activity that ensures a continuous supply of goods and services while mitigating risks through contract negotiation, cost modeling, quality control, and other critical supply factors (Lysons & Farrington, 2020, p. 4). It aims to deliver customer value by minimizing total cost of ownership, which includes not only the purchase price but also factors like transportation, storage, and quality (Christopher, 2016, pp. 29-30).

Modern procurement integrates digital technologies such as artificial intelligence (AI) and blockchain to enhance transparency, efficiency and productivity (Charles et al., 2023). It also aligns with corporate social responsibility (CSR) objectives, promoting ethical sourcing and sustainability (Quarshie et al., 2016).

The purchasing function focuses on the transactional aspects of procurement, such as ordering and receiving goods. It plays a critical role in ensuring that organizations acquire the right materials at the right time and cost (Benton, 2020, pp. 89-92).

Tactical purchasing includes determining needs, supplier selections, evaluating bids and contract negotiations. (Van Weele & Rozemeijer, 2022, pp. 36-40). Order function handles routine buying decisions, such as issuing purchase orders, expediting, communication and evaluation (Van Weele & Rozemeijer, 2022, pp. 42-45).

The purchasing process is a structured approach that organizations use to acquire goods and services to support their operations. A widely recognized framework for this process is the linear procurement process model, which outlines the sequential steps in procurement from need identification to supplier evaluation (Van Weele & Rozemeijer, 2022, p. 7). This model ensures cost-effectiveness, quality control, and supplier reliability.

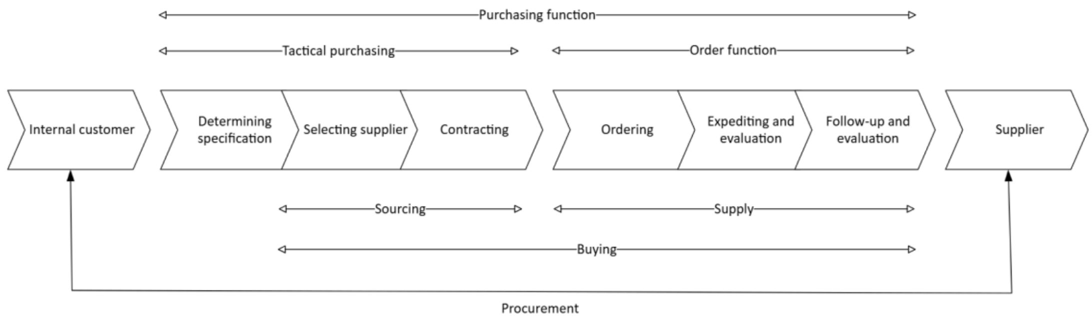


Figure 1. The linear procurement process model and some related concepts based on Van Weele & Rozemeijer (2022, p. 7).

Van Weele and Rozemeijer (2022, p. 31) describe the linear procurement process model as consisting of six key stages. Specification Phase – The process begins with identifying the need for a product or service. This phase involves defining specifications based on quality, quantity, and technical requirements (Van Weele & Rozemeijer, 2022, pp. 35-36).

Supplier Selection and Assessment – Procurement teams evaluate potential suppliers based on cost, delivery time, quality assurance, financial stability, reliability and capacity. This may involve market research, utilizing pre-existing supplier databases, or issuing requests for information (RFIs) or requests for proposals (RFPs) (Van Weele & Rozemeijer, 2022, pp. 36-38)

Negotiation and Contracting – After selecting a supplier, terms and conditions, including price, delivery schedules, and payment terms, are negotiated and formalized in a contract (Van Weele & Rozemeijer, 2022, pp. 38-42).

Ordering Process – The organization places purchase orders with the selected supplier, often integrating digital procurement systems for efficiency. In some cases, the contract acts as a purchase order, in other cases purchase order is placed based on frame agreement (Van Weele & Rozemeijer, 2022, pp. 42-43).

Expediting Process – Ensuring timely delivery and quality. This phase involves tracking orders, resolving delays, and maintaining supplier communication. Quality control might require inspections to the supplier's site acceptance tests (Van Weele & Rozemeijer, 2022, pp. 43-44).

Follow-up and Evaluation – Includes handling warranty claims, penalties, excess/minor work and archiving procurement files. Organizations assess supplier performance through key performance indicators (KPIs), ensuring compliance with contractual obligations and identifying areas for improvement (Van Weele & Rozemeijer, 2022, p. 44).

In the Figure 1 Van Weele & Rozemeijer (2022, p. 7) state that selecting supplier and contracting parts are included in the concept of sourcing, which includes identifying, evaluating, and selecting suppliers to meet an organization's purchasing needs (Van

Weele & Rozemeijer, 2022, p. 98). Effective sourcing strategies enable firms to optimize costs, improve supplier relationships, and enhance supply chain resilience.

The three later parts in the figure 1. Van Weele & Rozemeijer (2022, p. 7) are included in the concept of supply, which in this case refers to flow of goods and services from suppliers to customers after the supplier is selected.

Buying differs from the broader concept of procurement in that sense that determining specifications is not included the concept, therefore it is focusing more purely on the commercial activities (Van Weele & Rozemeijer, 2022, p. 8).

2.1.1 Services Purchasing

The purchasing of services differs significantly from purchasing tangible goods due to the intangibility, variability, and perishability of services. Organizations must adopt specialized procurement strategies to ensure the cost-effectiveness and quality of services procured (Van Weele & Rozemeijer, 2022, pp. 127-132). Effective service procurement requires clear specifications, performance-based contracts, and strong supplier relationship management. Figure 2 depicts the main characteristics of technical services procurement.

Figure 2. Characteristics of services purchasing.

Defining the scope of the work is important so that service provider knows what to deliver and accomplish. Axelsson and Wynstra (2002, p. 144) argue that it can be specified in a four different way: resources and capabilities; supplier processes needed to produce the service; functionality or performance of the service; or economic value for the customer.

The scope of work should clearly specify the nature of the services needed, including where, when, and to whom they will be provided. Additionally, the conditions under which the services are to be delivered must be outlined, along with the expected standards or performance levels. The agreement should also define the initial provision period and any renewal intervals. Furthermore, the purchaser's roles, if any, such as assistance with coordination, provision of equipment, staff support, or research, should be explicitly stated (Lysons & Farrington, 2020, p. 635).

Vendor selection begins with clearly defining the needed service, including specific performance indicators (KPIs) and desired outcomes. Market research and Requests for Information (RFIs) are then used to identify potential vendors and gather preliminary data on their capabilities. Shortlisted candidates collaborate with the buyer to refine service specifications, ensuring a mutual understanding of requirements. The evaluation process considers both quantifiable "hard" criteria like price, service quality, technical capabilities, and compliance, as well as qualitative "soft" criteria such as trust, cultural fit, innovation potential, and past performance. Crucially, the evaluation also assesses the vendor's ability and willingness to collaborate effectively throughout the relationship (Van der Valk & Rozemeijer, 2009).

Effective contract management ensures that agreements are clearly defined, monitored, and enforced, minimizing risks, controlling costs, and maximizing value in purchasing. A well-structured contract should include key elements such as a scope of work outlining deliverables and timelines, performance metrics to measure success, roles and responsibilities of both parties, payment terms specifying costs and invoicing procedures, penalty and incentive clauses to enforce compliance, dispute resolution mechanisms for conflict management, confidentiality and security provisions to protect sensitive information, and termination clauses detailing exit strategies and conditions for contract dissolution (Benton, 2020, pp. 435-443).

Cost and value analysis in service purchasing has shifted from a traditional focus on cost reduction to a broader emphasis on value creation. Modern procurement strategies assess suppliers not only on price but also on factors such as quality, innovation, sustainability, and alignment with business objectives. Performance-Based Contracts (PBCs) are increasingly used to tie supplier payments to measurable outcomes, ensuring accountability and service excellence. Additionally, advancements in digital tools like big data analytics and AI enable better cost-value assessments, optimizing procurement decisions. Ultimately, businesses must balance cost efficiency with long-term strategic value to gain a competitive advantage in service purchasing (Hofmann et al., 2020).

Risk management is crucial in services purchasing. Increasing complexity raises management challenges and costs, making coordination essential for effectively integrating external services, while power imbalances can lead to opportunism or inefficiencies, and dependence influences negotiation leverage and long-term stability (Benton, 2020, pp. 431-434). According to Kraljic (1983) risk management in procurement requires a proactive approach to identify and mitigate potential challenges.

Effective services procurement can lead to cost savings and increased efficiency. By outsourcing non-core activities, organizations can focus on their primary business functions and leverage the expertise of specialized service providers (Heinis et al., 2022).

2.1.2 Process in Industrial Management

Process theory in industrial management is rooted in operations research, quality management, and systems thinking. Taylor (1911) introduced scientific management, advocating for systematic process optimization through time and motion studies. His work laid the foundation for structured industrial processes focusing on efficiency and productivity.

The Input–Process–Output (IPO) model has its roots in systems theory, which views organizations as dynamic entities that process inputs into outputs through a series of interrelated steps (Bertalanffy, 1968). The model is particularly influential in operations research and industrial engineering, where it helps to analyze and improve production efficiency.

Deming's book "Out of the Crisis" in 1986 emphasized continuous process improvement through statistical quality control and The Shewhart Cycle today known as Plan-Do-Check-Act (PDCA) cycle (Deming, 2000, p. 88). His contributions significantly influenced Total Quality Management (TQM) and Lean Manufacturing principles, both of which aim to minimize waste and improve process efficiency.

Process theory is widely applied in industrial management through frameworks such as Lean Manufacturing, Six Sigma, and Business Process Reengineering (BPR). Lean Manufacturing, developed from the Toyota Production System, focuses on eliminating non-value-adding activities to enhance efficiency and reduce costs (Ohno, 1988).

Six Sigma, introduced by Motorola, combines process efficiency with data-driven decision-making to reduce defects and variations in production (Pyzdek & Keller, 2014). This methodology has been widely implemented in manufacturing, healthcare, and service industries to improve quality and operational performance.

BPR, introduced by Hammer and Champy (1993, pp. 90-91), advocates for radical process redesign to achieve substantial performance improvements. By reengineering workflows and leveraging technological advancements, organizations can enhance productivity and competitiveness.

With advancements in Industry 4.0, digital transformation has become integral to industrial process management. Davenport and Short (1990) highlighted the role of Information Technology (IT) in business process redesign, emphasizing how automation optimizes workflows and enhances decision-making.

Robotic Process Automation (RPA) and Artificial Intelligence (AI) have revolutionized industrial management by automating repetitive tasks and improving efficiency. Van der Aalst (2012) introduced process mining as a data-driven approach to analyze and optimize industrial workflows.

2.2 Artificial Intelligence

Artificial Intelligence (AI) refers to the capability of machines to perform tasks that typically require human intelligence, including learning, reasoning, problem-solving, perception, and language understanding (Russell & Norvig, 2022, pp. 19-23).

Machine learning (ML) is a subset of artificial intelligence that enables computers to identify patterns and make data-driven decisions without being explicitly programmed. It encompasses a variety of techniques, such as supervised learning (using labeled datasets) and unsupervised learning (finding patterns in unlabeled data) (Le et al., 2020).

Deep learning is a subset of machine learning that utilizes multi-layered artificial neural networks to automatically learn hierarchical features from data. Unlike traditional machine learning algorithms that rely on manual feature engineering, deep learning models, especially convolutional neural networks (CNNs), are capable of learning directly from raw data such as images, text, or audio (Alzubaidi et al., 2021).

Natural Language Processing (NLP) is a subfield of artificial intelligence that focuses on enabling computers to understand, interpret, and generate human language.

Contemporary NLP heavily leverages machine learning and deep learning techniques to process and analyze vast amounts of text and speech data (Rayhan, 2024).

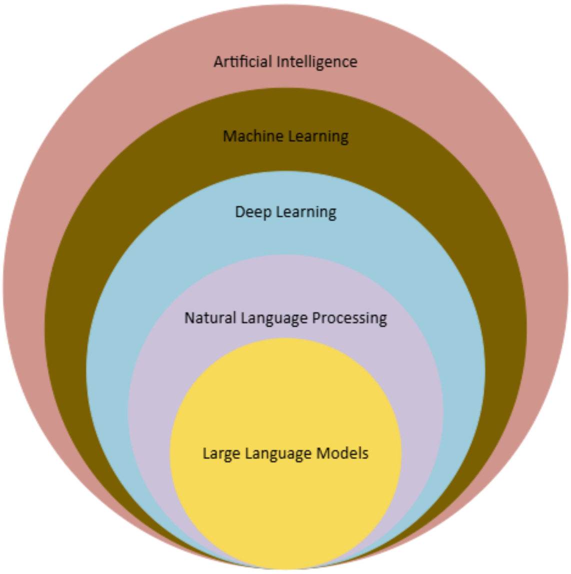


Figure 3. AI and related concepts including LLMs.

In the Figure 3 is illustrated the AI and related concepts including LLMs, chapter 2.2.2 introduces the concept of Large Language Models.

2.2.1 History and Development of AI

The formal study of Artificial Intelligence (AI) began in the 20th century with the development of computational theory. Alan Turing (1950) proposed the Turing Test as a measure of machine intelligence, arguing that a machine capable of human-like conversation could be considered intelligent. The concept of symbolic reasoning and logic programming was later advanced by McCarthy, who coined the term "Artificial Intelligence" in 1956 (McCarthy et al., 2006, p. 12).

Phases of AI Development

The Birth of AI (1950s-1960s): The Dartmouth Conference in 1956 marked the official founding of AI as a field of study. Early programs such as the Logic Theorist and General Problem Solver demonstrated the potential of symbolic reasoning (Newell & Simon, 1956)

AI Winters (1970s-1980s): Funding and interest in AI declined due to unfulfilled promises and technical limitations. During this period, expert systems gained traction in specific domains such as medical diagnostics (Feigenbaum, 1984).

The Rise of Machine Learning (1990s-2000s): Advances in statistical methods and increased computing power revitalized AI. Neural networks, which had been largely abandoned, reemerged with improved algorithms and greater computational capacity (LeCun et al., 2015)

Deep Learning and Modern AI (2010s-Present): The development of deep learning and big data analytics revolutionized AI applications, from image recognition to natural language processing. Breakthroughs such as AlphaGo demonstrated the power of AI in complex decision-making tasks (Silver et al., 2016).

AI has transitioned from theoretical explorations to real-world applications, influencing various industries such as healthcare, finance, and autonomous systems (Brynjolfsson & McAfee, 2017). Ethical considerations, including bias, transparency, and accountability, remain critical areas of discussion as AI continues to evolve (Bostrom, 2014).

2.2.2 Large Language Models

Large Language Models (LLMs) are a subset of artificial intelligence (AI) designed to process and generate human-like text by leveraging vast amounts of data and sophisticated neural network architectures. These models utilize deep learning techniques, particularly transformer-based architectures, to understand, predict, and generate text in various contexts (Vaswani et al., 2017).

Core Principles of LLMs

Introduced by Vaswani et al. (2017), the transformer model revolutionized natural language processing by enabling parallel processing of text and significantly improving the efficient handling of long-range dependencies in language. This architecture, based on self-attention mechanisms, allows models to capture contextual relationships more effectively, leading to substantial advancements in machine translation, text generation, and various other Natural Language Processing (NLP) applications.

The Self-Attention mechanism enables LLMs to dynamically focus on different parts of the input text while generating responses. By assigning varying levels of importance to different words or phrases, the model can better capture contextual relationships and dependencies within the text. This ability significantly enhances the model's understanding of long-range dependencies and coherence in natural language generation, leading to more accurate and contextually relevant responses (Brown et al., 2020).

LLMs are typically pre-trained on vast datasets containing diverse textual information and then fine-tuned for specific applications to enhance their performance in specialized domains. These applications can range from medical diagnostics, where they assist in analyzing patient data and generating reports, to legal text generation, where they help draft contracts and summarize case laws, and conversational AI, where they enable more natural and context-aware interactions with users (Radford et al., 2019).

Technologies in LLMs

Deep neural networks, particularly transformers, form the backbone of LLMs, enabling high-performance language modeling by capturing long-range dependencies and contextual relationships within text (Devlin et al., 2019). These architectures allow LLMs to process vast amounts of textual data efficiently, improving their ability to understand and generate human-like text.

To facilitate efficient processing and the generation of coherent text sequences, text is broken down into smaller units known as tokens. Tokenization techniques, such as byte pair encoding (BPE) and WordPiece, help LLMs handle diverse vocabulary and complex linguistic structures while maintaining contextual integrity (Sennrich et al., 2016).

Reinforcement Learning with Human Feedback (RLHF) is a critical technique used in models like ChatGPT, where AI-generated responses are refined based on human preferences. This iterative training approach improves the model's ability to generate contextually relevant and user-aligned responses by incorporating feedback from human evaluators, ultimately enhancing its effectiveness in real-world applications (Christiano et al., 2017, p. 4765).

Ethical Considerations in LLMs

Language models are trained on large datasets, which often contain biases related to factors like gender, race, and socioeconomic status. These biases can be reflected and even amplified by the models, leading to ethical concerns when used in decision-making processes. For example, biased models could perpetuate inequalities in areas like hiring or law enforcement. Bender et al. (2021) emphasizes the need to address these biases to ensure ethical use of AI systems.

LLMs, while powerful and versatile, can generate incorrect or misleading information due to their reliance on patterns found in vast amounts of data. These models do not have a true understanding of the content they produce, and their responses can sometimes be factually inaccurate, contradictory, or biased. This raises significant concerns about their reliability, especially when used in contexts that require factual accuracy, such as in healthcare, law, or journalism. Zellers et al. (2019) highlight these concerns, pointing out the potential risks of LLMs inadvertently spreading misinformation or being misused in ways that could harm decision-making processes.

The use of large-scale data for training foundation models poses risks related to user privacy and data security. This data is primarily sourced from the internet, including websites, social media, and other publicly available information. Foundation models are trained on massive amounts of data, and the models can sometimes reproduce this data verbatim, which could violate the privacy of the people who created the data or the copyrights of the owners of the data. Additionally, the models can sometimes leak sensitive information about individuals or organizations, which could be used for malicious purposes (Bommasani et al., 2022).

Lee et al. (2025) found that AI can reduce the perceived effort of critical thinking, leading to over-reliance on AI and decreased independent problem-solving. Their survey of knowledge workers revealed that while AI assistance can make tasks seem easier, especially for those confident in AI's abilities, this can result in decreased critical thinking and a tendency to accept AI outputs without scrutiny. The authors warn that this over-reliance may hinder the development of essential critical thinking skills.

"In the Belly of AI" (Poulain, 2024) exposes the exploitative conditions faced by AI trainers, the hidden workforce powering artificial intelligence. These workers, often in developing countries, perform essential tasks for low wages and under precarious employment. The film highlights the dehumanizing nature of the work and the industry's reliance on this exploitation, connecting it to environmental damage from data centers. It implicitly critiques "longtermism" suggesting that the pursuit of future AI benefits comes at the cost of present-day harms to both people and the planet.

2.2.3 Google Gemini

Google Gemini is a state-of-the-art artificial intelligence model developed by Google DeepMind, designed to advance the capabilities of large language models (LLMs) by integrating multimodal learning, deep reinforcement strategies, and real-time adaptability. Gemini represents a significant step forward in AI, incorporating the latest advancements in machine learning, natural language processing (NLP), and artificial general intelligence (AGI) research (Google, 2023).

Unlike conventional language models, Google Gemini demonstrates multimodal capabilities by integrating text, images, audio, and video processing, enhancing its versatility in understanding and generating diverse types of data. The model processes and combines different data types seamlessly, enhancing its ability to understand and generate content in complex scenarios (Panchal et al., 2024).

Google Gemini leverages Reinforcement Learning from Human Feedback (RLHF) to fine-tune its models and better align them with human preferences. This involves training a reward model based on human feedback and then using reinforcement learning to optimize the model's responses based on this reward model. This iterative process leads to improved model performance and the generation of higher-quality outputs (Gemini Team, 2023).

2.3 Artificial Intelligence in Procurement

Artificial intelligence (AI) is rapidly transforming various aspects of businesses, and procurement is no exception. AI's ability to analyze vast amounts of data, automate tasks, and provide insights is revolutionizing how organizations source, procure, and manage goods and services. This part explores the applications, benefits, and challenges of using

AI in procurement as illustrated in the Figure 4, drawing upon academic research and industry insights of the field of study so far.

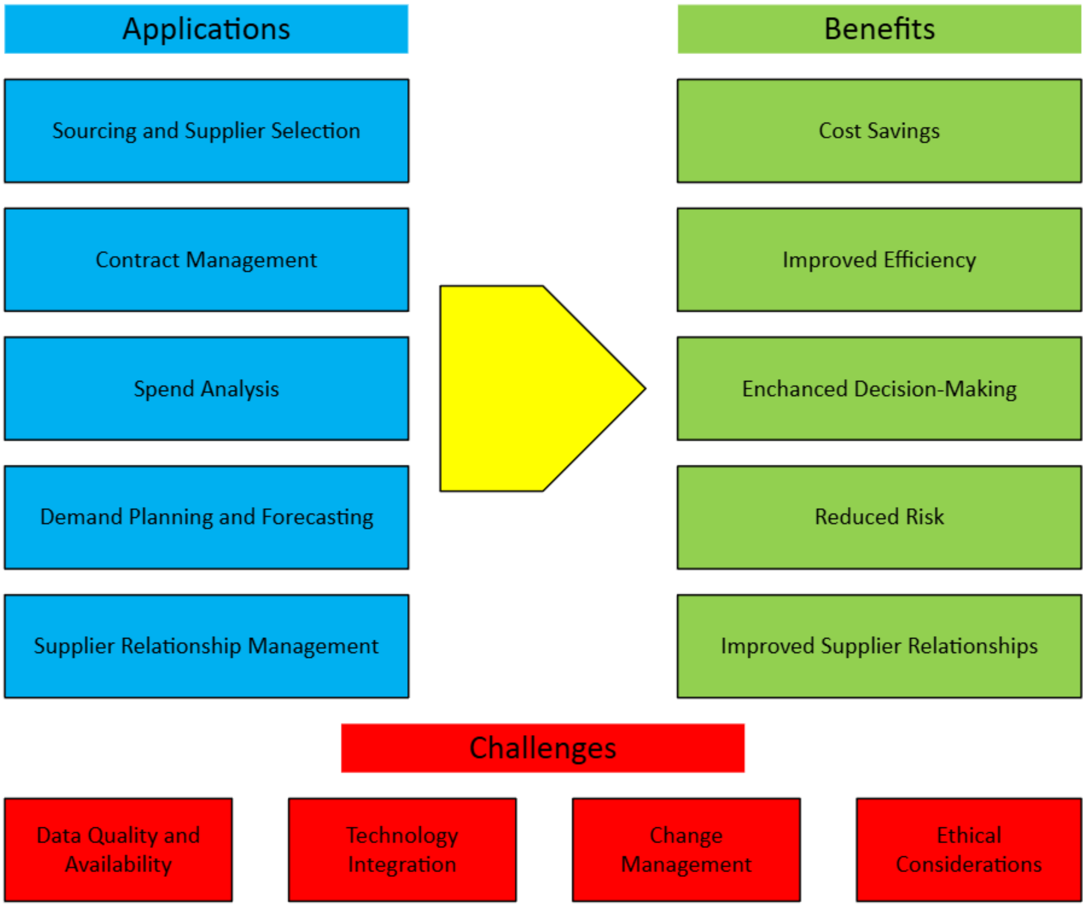


Figure 4. AI in Procurement.

AI is being applied across various stages of the procurement process, including:

Sourcing and Supplier Selection: AI can analyze supplier data, market trends, and risk factors to identify the most suitable suppliers (Guida et al., 2023). AI-powered platforms can automate the evaluation and onboarding of new suppliers, ensuring compliance and reducing risk (Spreitzenbarth et al., 2024)

Contract Management: AI can extract key information from contracts, track deadlines and obligations, and identify potential risks and opportunities (Toorajipour et al., 2021).

This automation frees up procurement professionals to focus on more strategic activities (Burger et al., 2023).

Spend Analysis: AI can analyze spending patterns, identify cost-saving opportunities, and provide insights into procurement performance (Allal-Cherif et al., 2021). This data-driven approach enables organizations to optimize their procurement strategies and improve their bottom line.

Demand Planning and Forecasting: AI can analyze historical data, market trends, and external factors to predict future demand for goods and services (Culot et al., 2024). This helps organizations optimize inventory levels, reduce stockouts, and improve supply chain efficiency.

Supplier Relationship Management: AI can monitor supplier performance, identify potential issues, and facilitate communication and collaboration (Wamba et al., 2024). This helps organizations build stronger relationships with their suppliers and improve overall supply chain resilience.

The adoption of AI in procurement offers several benefits, including:

Cost Savings: AI can automate tasks, optimize processes, and identify cost-saving opportunities, leading to significant cost reductions (Guida et al., 2023).

Improved Efficiency: AI can automate repetitive tasks, freeing up procurement professionals to focus on more strategic activities. This improves overall efficiency and productivity (Spreitzenbarth et al., 2024).

Enhanced Decision-Making: AI can analyze vast amounts of data and provide insights that enable better decision-making. This leads to more informed procurement strategies and improved outcomes (Toorajipour et al., 2021).

Reduced Risk: AI can identify potential risks and compliance issues, helping organizations mitigate risks and ensure ethical sourcing practices (Allal-Cherif et al., 2021).

Improved Supplier Relationships: AI can facilitate communication and collaboration with suppliers, leading to stronger relationships and improved supply chain resilience (Wamba et al., 2024).

Despite the numerous benefits, there are challenges associated with AI implementation in procurement:

Data Quality and Availability: AI algorithms require large amounts of high-quality data to function effectively. Organizations need to ensure data accuracy, completeness, and accessibility (Culot et al., 2024).

Technology Integration: Integrating AI tools with existing procurement systems can be complex and require significant investment. Organizations need to carefully evaluate and select AI solutions that align with their needs and infrastructure (Burger et al., 2023).

Change Management: Implementing AI requires changes to procurement processes and roles. Organizations need to effectively manage change and ensure buy-in from procurement professionals (Guida et al., 2023).

Ethical Considerations: AI algorithms can perpetuate biases present in the data they are trained on. Organizations need to be mindful of ethical considerations and ensure fairness and transparency in AI-driven procurement decisions (Spreitzenbarth et al., 2024).

2.4 Lean Six Sigma

Lean Six Sigma (LSS) is a methodology that integrates Lean principles, which focus on waste reduction, with Six Sigma strategies aimed at minimizing variation and improving quality. This approach has been widely adopted across various industries to enhance operational efficiency and customer satisfaction (Antony, 2016, pp. 34-35).

Lean Six Sigma integrates principles from Lean manufacturing and Six Sigma methodologies. Lean manufacturing, rooted in the Toyota Production System (TPS), emphasizes waste elimination, process efficiency, and continuous improvement (Antony et al., 2017). On the other hand, Six Sigma, developed by Motorola in the 1980s, is a data-driven approach that focuses on reducing process variation using the Define, Measure, Analyze, Improve, and Control (DMAIC) framework (Pyzdek & Keller, 2014, pp. 3-5). The combination of these methodologies aims to enhance quality and efficiency by leveraging the strengths of both approaches.

2.4.1 Lean

Lean was developed by Toyota to reduce inefficiencies and optimize production. It is built on five core principles: identifying value, mapping the value stream, creating flow, establishing pull, and striving for perfection. Value is defined from the customer's perspective. The goal is to determine what customers truly need and focus on activities that directly contribute to fulfilling those needs. Once value is defined, the next step is mapping the value stream, which involves analyzing every step in the process to see how value flows from raw materials to the final product (Liker, 2004, pp. 30-34).

Flow refers to the seamless movement of work through the production process without interruptions or delays (Liker, 2004, pp. 61-77). A pull system ensures that products are only produced when there is customer demand, preventing overproduction and reducing excess inventory (Liker, 2004, pp. 79-89). The final principle emphasizes continuous improvement (Kaizen). Lean is a never-ending journey, where every employee is encouraged to contribute small, incremental improvements daily (Liker, 2004, pp. 351-374).

The methodology encourages organizations to minimize non-value-added activities, known as the seven wastes (muda): overproduction, waiting, transport, extra processing, inventory, motion, and defects. Eliminating these non-value-added activities operating efficiency can be improved significantly (Ohno, 1988, pp. 17-18).

2.4.2 Six Sigma

Six Sigma is a data-driven methodology focused on improving process quality by identifying and eliminating defects. Originating from Motorola in the 1980s, Six Sigma employs statistical analysis to enhance operational efficiency and reduce variation in production and service processes (Pyzdek & Keller, 2014, pp. 3-4).

Six Sigma is built on two primary methodologies: Define, Measure, Analyze, Improve, and Control (DMAIC) for existing processes and Define, Measure, Analyze, Design, and Verify (DMADV) for new processes. These frameworks provide structured problem-solving techniques aimed at reducing process variability and enhancing performance. The methodology relies on statistical tools such as process capability analysis, regression analysis, and hypothesis testing (Deepali Kishor Desai, 2010, pp. 41-61).

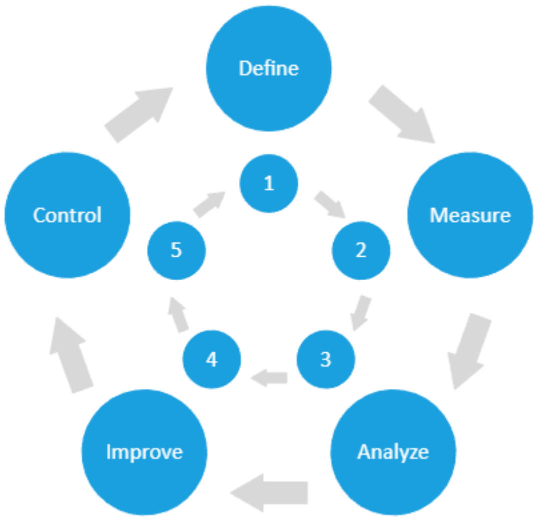


Figure 5. DMAIC based on Deepali Kishor Desai (2010, p. 43).

DMAIC consists of five steps as illustrated in the Figure 5.

The initial phase of Six Sigma, Define, focuses on clearly outlining the project's purpose. This involves defining the problem, setting goals, identifying customers and their needs, and creating a project plan. By asking "what, why, who, and how" questions, teams clarify the "big" issues, map processes, and develop a DMAIC charter. This process initiates business rethinking by establishing a solid foundation for improvement (Deepali Kishor Desai, 2010, pp. 43-44).

During the Measure phase, it is essential to determine the current performance of the service process by evaluating key metrics such as process yield, defects and capability indices. Identifying what to measure and establishing an appropriate measurement system are crucial steps. A benchmarking exercise helps assess how the process compares to industry standards or competitors. Finally, analyzing strengths and weaknesses allows for identifying gaps and areas for improvement (Antony, 2006, p. 240).

In the Analysis phase, the goal is to identify and address the root causes of defects within a process. This involves investigating sources of variability that contribute to defects and prioritizing them for deeper analysis. Understanding data patterns and distributions helps pinpoint key process variables that may be linked to performance issues. Additionally, evaluating the financial impact of improvements ensures that proposed solutions offer measurable benefits to the organization (Antony, 2006, p. 240).

The Improve phase in DMAIC focuses on implementing solutions discovered in the Define, Measure, and Analyze phases. Design of Experiments (DOE) is heavily utilized, systematically testing and optimizing solutions to ensure effectiveness. This stage aims to improve the process by implementing data-driven solutions and verifying their impact through experimentation (Breyfogle, 2003, pp. 549-666).

To ensure the improvements are rooted, the Control phase is essential. It can be thought as the "maintenance" stage where you keep the positive changes going. This involves a few key steps: First, control charts are used to continuously monitor the process and alert you to any deviations or changes. Second, detailed control plans are created to document the improved process, including the control mechanisms and how to respond to any issues that arise. Third, the concept of Poka-yoke, or error-proofing, is implemented to prevent defects and mistakes from happening in the first place. Finally, realistic tolerances are set to balance the desired quality with what the process can consistently achieve (Breyfogle, 2003, pp. 669-718).

2.4.3 Benefits and Challenges of Lean Six Sigma

One of the primary benefits of Lean Six Sigma is its ability to streamline processes and eliminate waste, leading to enhanced efficiency and productivity. By identifying and reducing non-value-added activities, organizations can optimize workflows and allocate resources more effectively (Womack & Jones, 1997).

Six Sigma principles focus on reducing process variation and minimizing defects, leading to improved product and service quality. Through statistical analysis and data-driven decision-making, LSS ensures that organizations consistently meet customer expectations and regulatory standards (Pyzdek & Keller, 2014, pp. 3-5).

By eliminating inefficiencies and improving process performance, Lean Six Sigma leads to significant cost savings. Organizations utilizing LSS have reported substantial reductions in operational costs due to lower defect rates, improved cost of goods sold, reduced cost of poor quality, improved cycle times, and equipment effectiveness (Basu, 2011, pp. 25-26).

Lean Six Sigma enhances customer satisfaction by ensuring consistent product and service quality. By focusing on customer-defined value and reducing process variability, LSS leads to shorter lead times, better product reliability, and improved service delivery (Breyfogle, 2003, pp. 52-57).

LSS fosters a culture of continuous improvement and empowers employees to take an active role in problem-solving. Organizations that adopt LSS often see increased employee engagement due to structured training programs, leadership support, and data-driven decision-making (Nonthaleerak & Hendry, 2008).

Lean Six Sigma provides organizations with robust analytical tools to make informed decisions. Statistical process control, root cause analysis, and predictive analytics allow businesses to respond proactively to emerging challenges and prevent defects before they occur (Antony, 2006).

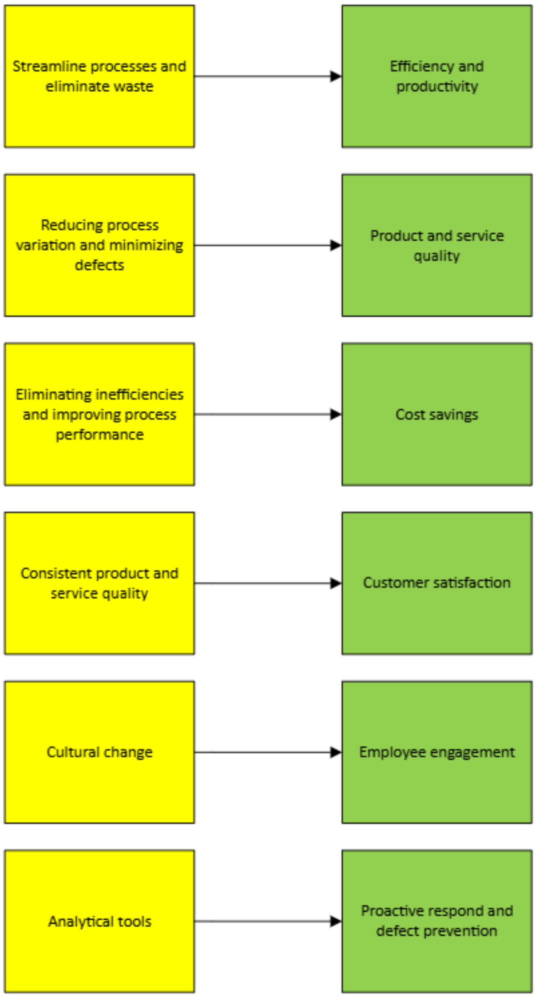


Figure 6. Benefits of LSS.

Figure 6 illustrates the causal relationship between Lean Six Sigma (LSS) implementation actions and resultant benefits.

One of the most significant barriers to Lean Six Sigma implementation is resistance to change. Employees and managers may be reluctant to adopt new processes due to fear of job displacement, unfamiliarity with LSS concepts, or skepticism about its effectiveness (Nonthaleerak & Hendry, 2008).

Although LSS is designed to improve cost efficiency, its initial implementation can be expensive. Organizations must invest in training, hiring certified professionals, and upgrading technology to support data-driven decision-making (Aljazzazen & Schmuck, 2022).

Lean Six Sigma (LSS) relies on sophisticated statistical tools, but many employees lack the necessary expertise to apply these tools effectively, leading to poor implementation and suboptimal results. Organizations must invest in training and educating their employees on these tools to ensure the success of LSS implementation. Additionally, organizations should also focus on hiring new employees skilled in emerging technologies (Samanta et al., 2024).

Successful LSS implementation requires strong leadership support. Without active involvement from top management, projects may lack direction, resources, and long-term sustainability. Once the initial project is completed, there is often a decline in momentum due to shifting business priorities, lack of reinforcement mechanisms, or failure to embed LSS principles into the company's culture (Breyfogle, 2003, pp. 24-27, 49).

Quantifying the return on investment (ROI) of LSS initiatives can be complex, particularly in service industries where benefits such as customer satisfaction and process efficiency are less tangible. Organizations must develop clear key performance indicators (KPIs) to accurately assess LSS impact (Antony, 2006).

Figure 7. Challenges of LSS.

Figure 7 illustrates the challenges associated with Lean Six Sigma implementation.

3 Method

This chapter details the methodology used to investigate how Google Gemini can help in the purchasing process. A design research approach as a research method including purchasing process analysis and pilot cases. Data is collected using observations, documentation and workshops. Lead time and defects are analyzed before and after Google Gemini's implementation.

3.1 Research Design and Approach

Research method of the study is adaption of design research method. Lean Six Sigma methodology is utilized to improve the process. Design science research model is illustrated on the Figure 8.

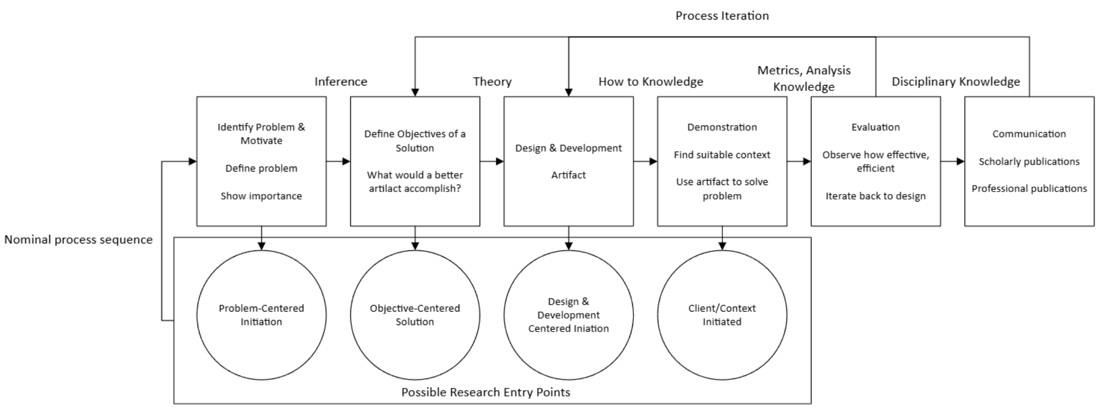


Figure 8. DSRM Process Model based on Peffers et al. (2007, p. 54).

The case company desires to find out how Google Gemini can help in the purchasing process and potentially it could bring significant improvements in efficiency and quality. Google Gemini frees up time by acting as an assistant through summarizing, analyzing, and more. Gemini also leads to improved productivity, allowing users to focus on strategic thinking, analysis, creative exploration, and boosting overall productivity. Furthermore, it contributes to well-being by lowering frustration and providing more time for thinking. Finally, Gemini results in improved quality by supporting the gathering of better information, providing more precise answers, enhancing content, fostering creative ideas, and generally improving quality.

Google Gemini offers clear benefits for tasks like email writing and information retrieval. Applying it to improve the case company's technical services purchasing process requires specialized knowledge of purchasing workflows, AI, and process improvement. Identifying potential applications within a specific company involves analyzing waste and defects in their current technical services purchasing process, and Gemini's capabilities. User input is gathered through workshops. Pilot projects are then defined, tested, measured and analyzed. All of this is something that a regular procurement professional doesn't have time to do during the daily work.

Purchasing Process Analysis

The purchasing process analysis aims to map all steps in the case company's technical services purchasing process, enabling the targeted discovery of Google Gemini's capabilities and their potential to reduce waste and defects. Detailed process mapping facilitates precise identification of waste and defects, while targeted procurement workshops gather data to pinpoint specific areas for Google Gemini deployment.

Potential pilot cases are evaluated and scored using the Pugh matrix (Pugh, 1991) based on five criteria: Time saving, Complexity to use, Scalability to other cases, Fit for Gemini, and Quality improvement of output. Time saving refers to the potential efficiency improvement of the process step. Specifically, does implementing Google Gemini reduce the end user's time within that step? Complexity to use assesses the ease with which end users can deploy and benefit from Google Gemini within the process step. Scalability to other cases refers to the ease with which the same Google Gemini deployment can be applied to other process steps or purchasing processes. Fit for Gemini assesses the solution's suitability for Google Gemini within the process step. Quality improvement of output assesses whether the solution enhances the output quality of the process step. Based on the score three to four most potential cases are then selected as the pilot cases.

Pilot Cases

When the most potential cases are selected, they are designed meaning that the prompt text that is going to be given to the Google Gemini is planned and which kind of attachments or base information it would require Google Gemini to deliver desired outcome. When prompt including attachments is designed it can be run and based on the outcome the prompt needs to be modified or not. At the point where the outcome can't be improved anymore, the pilot cases are measured and analyzed. To assess Google Gemini's impact, each pilot case included five runs with both the old and new designs, measuring lead time and defects before and after implementation for comparative analysis.

3.2 Data Collection

Data is collected through observations, documentation and workshops enhance understanding of the current purchasing process and evaluate Google Gemini's capabilities, observations and documentation are used to identify potential inefficiencies and defects that Google Gemini could address. In the workshop, the technical services purchasing team assesses the existing process using a Customer Journey Map (Rosenbaum et al., 2017). At each step, they rate their emotional response on a scale of 0 to 5 (ranging from "Very Bad" to "Very Nice") and provide feedback for improvement.

4 Results

This chapter presents the results of the purchasing process analysis and the pilot cases.

4.1 Purchasing Process Analysis

This section defines the current purchasing process, presents Google Gemini's capabilities, showcases workshop results, and details the identification and selection of potential pilot cases.

4.1.1 Defining the Current Purchasing Process

To be able to find suitable steps in the process to utilize Google Gemini I needed to define the current technical services purchasing process. I used my own experience on the process, case company documentation such as instructions and expertise of the team leader of the technical services purchasing team. Figure 9 illustrates the current purchasing process of the technical services in the case company.

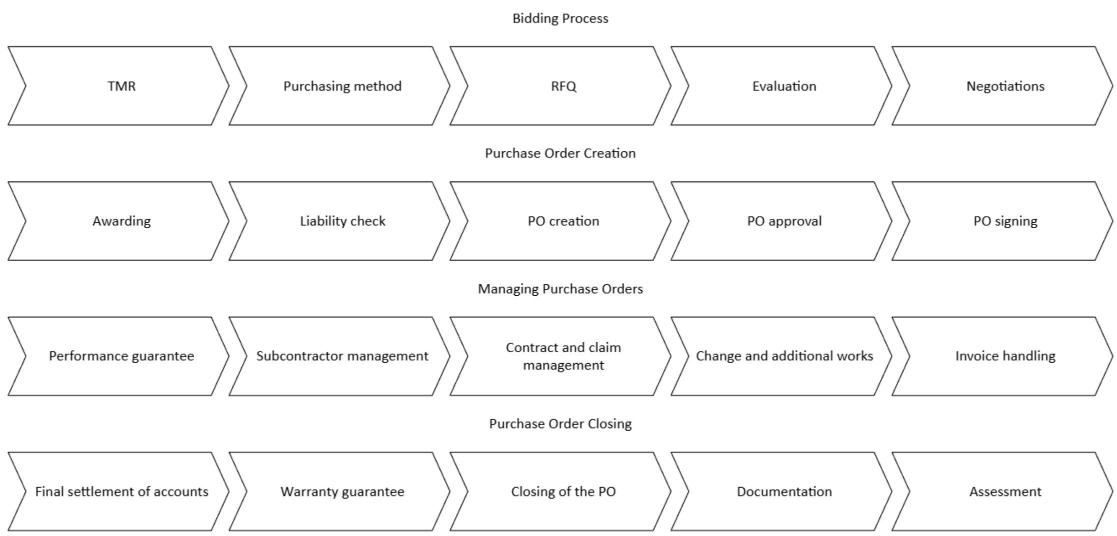


Figure 9. The current purchasing process.

Bidding Process

TMR (Technical Material Request): The procurement process begins with a TMR. The design engineer or CM (Construction Manager) creates the TMR in the Project Management system, based on the contracting plan by Project Procurement Manager (PPM). The PM (Project Manager) approves or rejects the TMR in the Project Management system. Approved TMRs are visible in the Buyer's Queue of the Project Management system.

Purchasing Method: The purchasing method (competitive bidding, single sourcing, or frame agreement) is determined by Project Procurement Manager (PPM) and the purchaser, based on the contracting plan and discussions with the project team. Case Company's procurement principle favors competitive bidding. Bidding is required for values over certain monetary value. Single sourcing (purchasing without bidding) requires permission and is used when a frame contract is unavailable or inapplicable. It is typically used when there is only one contractor for a service, or a specific contractor is preferred. Approval is always required for single sourcing. Competitive bidding is usually conducted when the scope is clearly defined and estimated monetary value is high enough so that time spent on the tender process will be worth it as competitive bidding is usually the way to get the best prices from the contractors. Frame agreements are used for smaller values, unclear scope, or unclear schedules. Contractors must be approved and audited by Case Company. For mechanical works, a technical audit is essential, and the contractor must have passed it.

RFQ (Request for Quotation): The purchaser creates a draft RFQ in the Project Management system and finalizes it in the Tendering system. Contractors are invited to submit quotations in the Tendering system. Several documents must be received or created off-system and uploaded to the Tendering system, including: General Terms (can be found from Google Drive), Technical documents (from engineering), Contract Programme (from CM), Limits of Responsibility and Costs (from CM), Job definition(s), drawing lists, isometric lists, isometrics, drawings, models, specifications (from engineering), HSEQ-instructions (can be found from Google Drive), other relevant and RFQ specific instructions and material, Unit price list (from engineering), Instalment table (Purchaser) and Turnaround specific instructions. The Tendering system is used to add vendors, set bid deadlines, and include questions for contractors. Purchasers often reuse questions from previous sourcing events.

Evaluation: The purchaser performs the commercial evaluation of the bids using template in Google Drive. The evaluation includes making the prices comparable in view of the scope of delivery and other factors influencing the price. The evaluation also includes estimated additional work performed on unit rates, with quantities prepared by Engineering/CM. CM conducts the technical evaluation of tenders.

Negotiations: If necessary, contract negotiations can be held with one or more contractors. The aim of negotiations is to ensure that the contractor has correctly understood the scope of the work, is familiar with the working conditions, and knows the safety requirements to be followed at work. The project team will support and participate in the negotiation process. At least the following members from the project team might be needed to participate: PM, CM, Site Supervisor and Design Engineer. Meeting minutes will be prepared for the contract negotiations, where all essential issues agreed in the negotiations are recorded, especially those where deviation from or addition to the invitation to bid documentation was made. The minutes of the meeting will be appended to the contract.

Purchase Order Creation

Awarding: The project team evaluates the suppliers based on the quotations, negotiations, and possible revised quotations. The purchaser makes a purchase recommendation/awarding in the Project Management system, and the Project Manager approves it in the Project Management system.

Liability Check: The purchaser creates an assignment in the Contractor Management system. The Contractor Management system verifies the contractor's compliance with contractor liability laws and Case Company's requirements. The Contractor Management system informs the purchaser that everything is ok after they have received and inspected the documentation provided by the contractor. Contractors and their subcontractors add their personnel working at the site to the Contractor Management system. Approved personnel receive site access via the Contractor Management system.

PO (Purchase Order) Creation: The purchaser creates PO draft in the Project Management system. The PO integrates with the ERP (Enterprise Resource Planning) system, where it is finalized, and details are added including text to the purchase order.

PO Approval: PO is approved in the ERP system depending on the approval limits.
The PO can be printed out from the ERP system to PDF format when PO is approved. Approvers are maintained in the ERP system and every project has different dedicated persons for approving.

PO Signing: When the value of the contract is over certain amount and the contract is made in writing (or in electronic form) and signed, it shall be signed by two persons unless the person has individually the right to sign on behalf of the Case Company entity.
The contractor has persons that sign the contract on behalf of the contractor. If the PO value is small, the contractor might give order confirmation via email.

Managing Purchase Orders

Performance Guarantee: Performance guarantee is defined in contractual terms of PO.

Approval of the first payment Report of the work/Invoice shouldn't be done before guarantee is received. Guarantee proof will be sent to AP (Accounts Payable) team and information will be added by responsible person to the Google Sheet.

Subcontractor Management: When the contractor adds subcontractors to the assignment on the Contractor Management system, personnel representing the Contractor Management system check the documents, after that the CM reviews the subcontractor, and the purchaser approves it.

Contract and Claim Management: If the contractor does not perform according to the contractual terms for instance delivery, quality, or safety. The purchaser prepares a claim to be sent for the contractor. There are separate instructions and templates for it and claims are marked down to google sheet.

Change and Additional Works: The budget of the PO needs to be kept up to date, when the project team identifies that scope of the work increases or other additional costs occur, the budget on the PO needs to be updated. CM needs to update the TMR in the Project Management system, PM approves changes, and Purchaser takes changes to the existing PO. The changes are also integrated to the ERP system. The contractor needs to fill additional works template and additional works need to be approved before they are started on the site. When work is performed actual hours/units need to be reported by the contractor and approved via Report of the work in the ERP system. Change order is applied when contractual terms of the PO are changed. New purchase order is printed out with changes.

Invoice Handling: In basic scenario, invoices are auto matched to the PO if the invoice has a PO and Report of the work number. However, if the invoice is for some reason not matching to any lines of the PO in the ERP system or there is more information needed, it comes to the invoice management workplace in the ERP system, where Purchaser can give instructions to the AP team.

Purchase Order Closing

Final Settlement of Accounts: CM will notify the Purchaser regarding completion of the work for the Purchase Order. The purchaser will create a folder to the Google Drive for the project close out, prepare a draft for the final settlement of accounts and will notify CM to fill in the first page. The purchaser will fill in the second page with required invoice information from the ERP system's PO. Once draft document is prepared, the Purchaser will invite CM, Site Supervisor and Contractor to a meeting. Upon mutual agreement, parties will sign the document and the Purchaser will enclose it to the ERP system's PO before closing the PO in the ERP system and archive it to the Document Management system. It is very important to have confirmation from the Contractor that all the work has been invoiced to avoid any open invoices in the future. There is a Google sheet template for the Final settlement of accounts.

Warranty Guarantee: Warranty guarantee is defined in the contractual terms of the PO. Approval of the last Report of the work/Invoice shouldn't be done before guarantee is received. Guarantee proof will be sent to the AP team and information will be added by responsible person to the Google sheet document.

Closing of the PO: PO is closed in the ERP system and the Project Management system to prevent further activities.

Documentation: All the project documentation will be archived in the Document Management system where document cards are opened for each PO and then information is added to the cards.

Assessment: Contractor performance is assessed, and lessons learned gathered. The contractor and project team fills in the HSEQ evaluation template. Sometimes for instance in turnarounds more comprehensive analysis on the contractor performance is conducted.

4.1.2 Google Gemini Capabilities

Google Gemini offers a range of capabilities including prompting, improving writing, organizing data, creating original images, summarizing information and surfacing insights, fostering meaningful connections with colleagues, researching unfamiliar topics, and spotting trends, synthesizing information, and identifying business opportunities.

Its features include "Help me write" in Google Docs and Gmail, "Help me organize" in Google Sheets, "Create image with Gemini" in Google Slides, creating background images in Google Meet, the ability to provide documents for Gemini to answer questions from in Google Drive, finding documents in Google Drive, comparing or summarizing documents in Google Drive, creating meeting minutes in Google Meet, utilizing Gemini in Google Chat, and creating videos with Google Vids.

Gemini on gemini.google.com offers several prompt engines tailored to different needs: 2.0 Flash for everyday tasks, the experimental 2.5 Flash with advanced reasoning, the experimental 2.5 Pro for complex tasks, and Deep Research powered by 2.5 Pro for in-depth research reports.

Canvas Mode is an interactive workspace within Gemini. After generating a first draft in Canvas, you can rapidly refine and ask Gemini for feedback to suggest edits. You can update specific sections or the whole draft and use the quick editor tools to change the tone, length, or formatting.

Google NotebookLM Plus is AI-powered research and writing tool where you upload documents and Gemini helps you understand, summarize, and generate ideas based on those sources. Helping you understand and create content based only on what you have given.

Google Gemini also offers "Gems," which are pre-made or custom prompts. Premade gems include a learning coach to break down complex topics, a brainstormer for inspiration, a career guide for professional development, a writing editor for feedback, and a coding partner for programming assistance. Users can also create custom gems to easily repeat specific prompts.

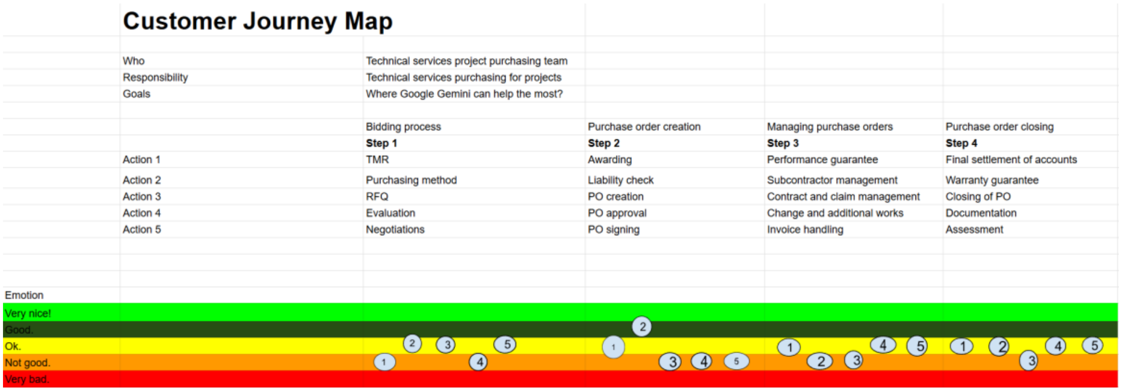
You can think of a prompt as starting a conversation with your AI-powered assistant, and like any conversation, you might use several prompts as it progresses. To write an effective prompt, there are four main areas to consider: persona, task, context, and format. While you don’t need to use all four, using a few will help. Use natural language as if you were speaking to another person. Be specific and iterate by telling Gemini for Workspace what you need it to do. Be concise and avoid complexity. Make it a conversation by fine-tuning your prompts if the results don’t meet your expectations.

Google Gemini's capabilities could potentially assist a technical services purchaser in their daily work. It could help create RFQ text, PO text, and RFQ documentation lists. It could prepare minutes of meetings and take notes. It could aid in bid evaluation, prepare for negotiations, and create awarding presentations. It could supervise vendor performance according to contractual terms and help in the final settlement of accounts.

A significant limitation of Google Gemini at this time is its inability to access our internal systems (such as ERP, Project Management, Contractor Management, and Document Management), although this capability may be developed in the future.

4.1.3 Workshop

At first, I only planned one workshop session of one hour for the Customer Journey Map on Tuesday 28th of January, eventually there were so much discussion in the workshop sessions, that I had to facilitate three additional workshop sessions for Tuesday 4th of February, Tuesday 11th of February and 18th of February.



Picture 1. Customer Journey Map.

As illustrated in Picture 1, the Customer Journey Map outlines the technical services purchasing team's process, from bidding to purchase order closure, aiming to identify where Google Gemini can improve efficiency. It tracks the team's emotional experience at each stage, highlighting potential pain points with numbered circles. The map focuses on optimizing internal workflows and leveraging AI for better process efficiency and output.

The emotion chart indicates that process satisfaction among the technical services purchasing team is low, with most actions falling within the 'Ok' or 'Not good' zones, and an average score of 2.65 out of 5. Only the Liability check received a 'Good' rating.

During the workshop sessions, participants made many development comments. The comments highlight inefficiencies and inconsistencies across multiple systems (project management, tendering, ERP), causing duplicate work and hindering operations. Key issues include challenging system usage, inadequate documentation, poor data transfer, and a lack of clear responsibilities. Specific problem areas include the bidding process, purchase order management, and related processes, all suffering from slow manual processing, errors, and communication gaps. Notes capturing these comments have been processed and summarized in the Appendix 1.

4.1.4 Potential Pilot Cases

Analysis of the current purchasing process, Google Gemini's capabilities, and insights from workshop sessions have identified the following steps where Gemini can potentially reduce waste and defects.

Awarding Presentation

The Awarding Presentation summarizes the tendering process and the rationale for vendor selection for audiences like project management, steering groups, or senior executives. Currently, creating these Google Slides presentations involves manually copy-pasting information from RFQs, Evaluation Templates, and Negotiation Memos. This manual method is time-consuming and prone to errors. Google Gemini could potentially automate this by generating the presentation and importing the necessary data directly from these source documents, quality of the output could also potentially increase.

Contract Supervisor

In this proposed pilot product, Google Gemini would assist in monitoring compliance with contract terms. For instance, it could analyze contracts to extract and track key obligations, such as delivery dates. Gemini would then automatically flag potential violations of the terms and conditions by contractors. This addresses a current challenge, as manually tracking numerous contracts and identifying breaches is difficult and time-consuming for the project team. By automating this oversight, the pilot aims to streamline the claims process, help ensure we capture all applicable liquidated damages for breaches, and ultimately generate cost savings.

Contractor Performance & HSEQ History

Can Google Gemini access and provide contractor performance data to aid in vendor evaluation? This includes historical information on additional work invoicing, adherence to milestones and deadlines, contract breaches, and HSEQ (Health, Safety, Environment, Quality) performance based on assessment data. Using this information would improve contractor selection.

Negotiation Coach

As a negotiation coach, Google Gemini aids preparation by recommending tactics informed by supplier segmentation and identifying leverage by comparing offers to should-cost benchmarks. It also supports the review of terms and conditions and evaluates detailed price tables (unit/hourly) against should-cost data and alternative bids to provide a clear negotiation landscape.

PO Text

Currently, creating Purchase Orders (POs) in the ERP system requires purchasers to manually add specific text for the printout. This often involves copying text from previous POs and adapting it using information from tendering documents—a time-consuming process prone to errors from manual data transfer. Google Gemini could automate this by extracting the necessary information directly from the tendering documents and generating the required PO text, improving efficiency and accuracy.

Pre-fills Negotiation MoM

Pre-filling negotiation meeting minutes is crucial; it ensures all critical points are covered and saves valuable time during the negotiation itself. Google Gemini directly addresses this need by automatically creating and populating minute templates using data from RFQs and quotations, thus streamlining preparation and eliminating the errors common in manual processes.

RFQ Document List

Currently, purchasers must manually compile a list of required documents before issuing an RFQ. This process is time-consuming and prone to errors. Google Gemini could automate the identification and listing of these documents, saving time and improving accuracy.

RFQ Text

When an RFQ is issued, it must be accompanied by introductory text (similar to a cover letter) that specifies the requesting party, the requirements, and the conditions for the supplier. Currently, this text is often copied from a previous RFQ, requiring manual updates with the specific details from the current tender documents.

Savings Tips

This pilot case idea aligns directly with the company's current strategic focus on enhancing procurement savings. Given the increasing pressure to generate greater cost reductions, this pilot will investigate Google Gemini's potential to assist the purchasing process by identifying savings opportunities and suggesting actionable measures.

Service RFQ Evaluator

Within the technical services purchasing process, a specific template is utilized for the commercial and technical evaluation of bids. This pilot proposes leveraging Google Gemini to automatically populate this template by extracting key information from the RFQ, submitted bids, and negotiation documentation. Automating this data entry, which is currently a time-consuming manual process, would significantly improve efficiency. Furthermore, Gemini could potentially score the bids with greater speed and consistency than manual evaluation allows.

Final Settlement of Accounts

The final settlement of accounts template must incorporate details of both invoiced and pending payment installments, along with costs for any additional works. While Google

Gemini could potentially automate adding this information, a significant limitation currently is its inability to access the ERP system where this data typically resides.

Records Negotiation MoM

Google Gemini offers a feature that automatically takes notes during Google Meet meetings. Could this functionality be effectively applied to negotiation sessions to potentially aid purchasers?

4.1.5 Selection of Pilot Cases

The Pugh Matrix (Table 1.) evaluates twelve potential pilot cases for implementing Gemini AI within the purchasing process. Each case was scored on a 1-5 scale across five criteria: "Time saving," "Complexity of use (Inverted)," "Scalability to other cases," "Fit for Gemini," and "Quality improvement". An "Overall score" was calculated for each pilot, presumably by summing the individual criteria scores. This matrix serves as a decision-making tool to identify the most promising pilot cases based on their assessed potential and feasibility.

Table 1. The Pugh Matrix.

Pilot Case	Time saving	Complexity of use (Inverted)	Scalability to other cases	Fit for Gemini	Quality improvement	Overall score
RFQ text	4	4	4	5	4	21
PO text	4	4	4	5	4	21
Contractor perf.	3	3	4	4	4	18
Savings tips	3	3	3	4	4	17
Negotiation coach	3	2	3	4	4	16

Pre-fills MoM	4	3	4	5	4	20
Final settlement	4	1	2	1	4	12
Contract supervisor	4	2	3	4	4	17
RFQ doc list	4	5	4	5	4	22
RFQ evaluator	4	2	3	4	4	17
Records MoM	3	4	3	3	3	16
Awarding presentation	4	4	3	5	4	20

The overall scores reveal distinct tiers among the potential pilot cases:

Top Tier: The highest scores were achieved by "RFQ doc list" (22), "RFQ text" (21), and "PO text" (21). These pilots generally scored very highly on "Fit for Gemini" (5/5) and benefited from strong scores in low complexity ("RFQ doc list" scored 5/5 inverted) and other criteria like time saving and quality improvement. Also scoring well were "Pre-fills MoM" (20) and "Awarding presentation" (20), both showing high Gemini fit and potential time savings.

Mid-Tier: Several pilots clustered in the middle range with scores from 16 to 18, including "Contractor perf." (18), "Savings tips" (17), "Contract supervisor" (17), "RFQ evaluator" (17), "Negotiation coach" (16), and "Records MoM" (16). These often-showed good potential benefits but tended to have lower scores for complexity (indicating higher perceived complexity) compared to the top tier.

Lower Tier: "Final settlement" scored significantly lower than others (12). This was primarily driven by very low scores in "Fit for Gemini" (1/5) and the inverted "Complexity of use" (1/5), due to system access issues noted previously.

The analysis indicates a strong alignment between Gemini's perceived capabilities and tasks involving document/list generation, which generally appear less complex to implement compared to pilots requiring deeper process integration or analysis of data from potentially inaccessible systems

Based on the high overall scores in the Pugh Matrix analysis, the following three pilot cases have been selected for implementation:
RFQ doc list (Score: 22)
RFQ text (Score: 21)
PO text (Score: 21)

These pilots represent the most promising opportunities identified through the scoring process, offering a strong combination of benefits (time saving, quality improvement), high fit with Gemini's capabilities, and manageable complexity.

Two other high-scoring pilots, "Pre-fills MoM" (20) and "Awarding presentation" (20), were considered but ultimately not selected for the initial implementation phase for the following reasons:

Pre-fills MoM: Although scoring well, this pilot was deferred due to concerns about its potential complexity in practice. Drafting negotiation Minutes of Meeting (MoM) requires careful alignment with the purchaser's specific strategy and ideas on how information should be presented, and which points need highlighting to ensure negotiations proceed effectively. Automating this effectively while maintaining the necessary strategic nuance was deemed too complex for an initial pilot.

Awarding presentation: This pilot was not selected because creating such presentations is an infrequent task within the procurement process. While Gemini could potentially offer benefits, the limited frequency means the overall impact on providing constant aid and efficiency improvements to the day-to-day process would be minimal compared to the selected pilots.

4.2 Pilot Cases

This section defines, tests, measures, and analyzes the selected pilot cases.

4.2.1 Defining and Testing Cases

RFQ Document List

The initial approach involved providing Google Gemini with both a link to the RFQ document list template and a link to the directory containing the relevant RFQ documentation. However, it became evident that Google Gemini was unable to directly generate a Google Sheets file or populate the provided template and return it in its completed form. While the system could produce a list of documents for subsequent importation, essential components such as the header (containing the company logo, date, and basic information) and the footer (containing the company name, business identification, and registered address) were absent. Furthermore, preliminary testing revealed inconsistencies in document identification, including incomplete listings and the generation of erroneous documents with fabricated revision dates.

A subsequent approach involved uploading documents directly using the “+” function within the Google Gemini interface. This method revealed limitations in supported file types; zipped files and folders could not be processed. The system required documents to be in Google Docs, Google Sheets, or PDF format, necessitating the conversion of design system documents. Although this approach initially resulted in the inclusion of document names from the template, this issue was subsequently rectified. Nevertheless, the generated lists remained limited to document names, excluding the required revision numbers and dates as specified in the template. An additional limitation was the restriction of ten documents per upload prompt, which presented a significant challenge given the large volume of documents associated with the technical services RFQ

A further attempt was made using the Canvas function within Google Gemini. While this approach facilitated the inclusion of revision numbers and dates, the persistent inclusion of documents derived from the template, despite explicit instructions to the contrary, remained a problem. Trials using versions 2.5 Pro and 2.5 Flash did not yield any noticeable improvements in this regard and neither did the NotebookLM, that wasn't able to read the revision numbers or dates.

The subsequent strategy involved a descriptive approach, where the required elements of the list were specified in detail, and documents were uploaded individually using the "+" function, thereby bypassing the difficulties encountered with link-based access. Due to the upload limitations, a selection of the most pertinent documents was employed for this test. A detailed description of the desired list format was provided. I noticed Gemini sometimes took revision number from file name, if file name has some wrong indication to revision number it takes it from the file name and not from the document. This approach yielded the most satisfactory results with the 2.0 Flash model. The 2.5 Pro model demonstrated comparable performance, offering no significant enhancements. Notably, the 2.5 Flash model exhibited unexpected difficulties in extracting the necessary information from the provided PDF documents.

The final prompt design, which produced the most acceptable outcome with the 2.0 Flash model, was as follows:
"You are a purchaser and need to create a document list from all the RFQ documents in the tendering process. The list will be sent out to the vendors with all the other documents. Use the following documents to fill the document list:
Insert RFQ envelope

Insert Contract Programme

Insert Limits of responsibility and costs

Insert HSE requirements

Insert Terms and conditions

Insert Document list(s) of Job definition(s)

Create it in a table format that can be exported to google sheets. Divide the commercial documents, HSE documents and technical documents. Include Document number, document revision number (don’t insert this from the document name), document revision date and document name.”

PO Text

I experimented with Gemini to generate a purchase order, providing the PO template, vendor quotation, and RFQ documents. Initially, it struggled with key details: misinterpreting the estimated price as a fixed cost, removing the PDF invoice email, omitting contact information, and leaving template fields unfilled. Canvas mode and the 2.5 versions (Pro and Flash) didn't improve results, with the 2.5 versions failing to read the RFQ folder.

Using Gemini's NotebookLM function offered slight improvement, but issues persisted with price representation (no PO value or calculation method), incorrect assignment of the project manager to technical matters (should be construction manager), and failure to identify other contacts.

Shifting my approach, I provided the quotation and RFQ with only high-level area headers for the PO. This yielded better results, but Gemini still seemed to have trouble processing the folder.

My most promising attempt involved providing essential documents with “+” function without using the folder link: contract programme, liabilities and responsibilities, job definition, quotation, document list, terms and conditions, and negotiation minutes. I also included specific instructions: avoid using the quotation for scope of work and use the contract programme for contact details. There were in addition problem with the invoicing addresses part as there were no document now, where Gemini could pick the addresses, as a result I decided that invoicing addresses need to be included in the prompt. Despite this targeted approach, I encountered the 'context window exceeded' error, advising smaller files. This design felt like the most effective strategy I could devise.

Similar problems arose with both the with the template and folder design and without the template and folder designs using 2.5 Flash (Experimental), while 2.5 Pro (Experimental) couldn't read the folder in the template and folder design. Testing the without the template and folder with 2.5 Pro (Experimental didn't show significant improvements over 2.0 Flash.

Here's the most optimal prompt design I used with 2.0 Flash:
“You are a purchaser and need to create a purchase order, use the following documents:
Insert Quotation from the supplier
Insert Job Definition
Insert Limits of responsibility and costs
Insert Contract Programme
Insert Document list
Insert Terms and conditions
Insert Negotiation MoM (if available)

Include following areas in the purchase order:
1. SCOPE OF DELIVERY (Don’t refer to the quotation when defining the scope of delivery)
2. HSE AND CODE OF CONDUCT
3. PRICE AND INVOICES

4. TIME SCHEDULE

5. WARRANTY

6. PERFORMANCE SECURITY

7. INVOICING ADRESSES

Insert invoicing adresses

8. CONTACT INFORMATION (Use the details in the contract programme)

9. ANNEXES”

RFQ Text

This pilot case shares similarities with the RFQ document list and PO text generation efforts. In this instance, the goal was also to have Gemini create text based on a template and provided documents.

Initially, I provided Gemini with the template and RFQ documents via a folder link. However, it relied on the existing information within the template. Neither Gemini 2.5 Pro nor 2.5 Flash yielded better results.

Subsequently, I omitted the template and instead specified the areas to be addressed in the RFQ text. None of the tested models (2.0 Flash, 2.5 Flash, or 2.5 Pro) could access the folder. Following the successful approach used for PO text generation, I then provided Gemini with the key documents and the required areas for the RFQ text. This method proved effective, except for the contact information and pricing. I then instructed Gemini to extract this information from the contract program and it worked better. However, still it was missing email addresses, but I then included it in the prompt. Again, neither 2.5 Flash nor 2.5 Pro offered significant improvements.

I also experimented with NotebookLM. While it allows for a larger number of documents, it doesn't support zip files or formats other than PDF, and the output quality wasn't significantly enhanced.

The most effective prompt was achieved with Gemini 2.0 Flash:"
“You are a purchaser and need to send RFQ to vendors. You need to create a text for the RFQ. Use following documents to create the RFQ text:
Insert Terms and Conditions
Insert Job Definition
Insert Limits of responsibilities and costs
Insert Contract Programme
Insert Document List
Insert HSEQ requirements

Include following areas in the RFQ:
1. SCOPE OF WORK
2. HSEQ AND CODE OF CONDUCT
3. PRICING AND PAYMENT TERM (Pricing from the contract programme)
4. CONTACT INFORMATION (From the contract programme, include email addresses)
5. FINAL DATE TO PLACE THE BID
6. TERMS AND CONDITIONS
7. FINAL DATE TO CONFIRM THE BID”

4.2.2 Measuring and Analyzing Cases

Each pilot case was tested five times with the same input material for both the old and the new design, and the resulting data was measured and analyzed.

RFQ Document List

Table 2 presents a comparison of lead time (in minutes and seconds) and the quantity of defects observed across five runs for both the old and new designs. The data reveals that the new design has resulted in a longer average lead time of 6 minutes and 18 seconds, compared to 6 minutes and 5 seconds for the old design – an increase of 13 seconds or 3.56%. Furthermore, the average number of defects per run has risen from 1 with the old design to 2.8 with the new design.

Table 2. Lead time and defects for the RFQ document list.

	Lead time (minutes and seconds)		Defects (quantity)	
Run number	Old design	New design	Old design	New design
1	7 min 28 sec	6 min 56 sec	4	0
2	7 min 5 sec	5 min 47 sec	0	8
3	5 min 22 sec	6 min 43 sec	1	0
4	5 min 20 sec	5 min 34 sec	0	3
5	5 min 14 sec	6 min 31 sec	0	3

Surprisingly, the old design's lead time became shorter than the new design's after the initial runs. Although the first run with the old design involved some manual errors (incorrect header information and a wrong revision number in one document), the lead time subsequently improved. In the third run, one defect occurred: a revision date mistake. This suggests a learning curve, where I became more proficient at accurately completing the template.

The new design lead time showed inconsistent improvement across runs. The first run had no defects, but the second had numerous defects, including missing revision dates (only one document had one) and a revision number error. The third run, despite arriving as three separate tables that required consolidation, had no defects. The fourth run presented issues with a wrong revision number, missing revision date, and a missing document name. Finally, the last run had documents missing revision dates, revision numbers, or both.

A key advantage of the old design is its unlimited document capacity, though users can manually add documents in the new design as well. However, selecting numerous documents for the list is more time-consuming in the old design, even if this manual approach potentially yields higher quality. Consequently, the new design might, in this specific aspect, produce a lower quality output. Ultimately, the preferred method hinges on the company's operational preferences and minimum output standards.

Manual document list creation is inefficient and prone to errors. It diminishes user satisfaction and relies heavily on individual user accuracy and motivation, leading to inconsistent quality. There's often a trade-off: meticulous users like myself produce accurate results but take longer, while faster users may introduce more defects. The new design doesn't fully eliminate this, as the manual step of pasting tables into a template for headers and footers remains a potential source of error.

The new design is unreliable, producing inconsistent output with more defects than the old version, and it fails to improve the process step's lead time.

PO Text

Table 3 compares the old and new designs across 5 runs, evaluating lead time (in minutes and seconds) and defect quantity. The new design reduced the average lead time from 5 minutes 28 seconds to 3 minutes 37.2 seconds—an improvement of 1 minute 50.8 seconds (33.78%). However, average defects increased from 0 per run for the old design to 1.4 per run for the new design.

Table 3. Lead time and defects for the PO text.

	Lead time (minutes and seconds)		Defects (quantity)	
Run number	Old design	New design	Old design	New design
1	8 min 22 sec	5 min 27 sec	0	2
2	5 min 47 sec	3 min 41 sec	0	1
3	4 min 16 sec	3 min 34 sec	0	1
4	4 min 33 sec	2 min 49 sec	0	1
5	4 min 22 sec	2 min 35 sec	0	2

Similar to previous RFQ document list old design runs where lead times decreased over runs, the lead time for this PO text also improved progressively. However, this time I did not make any errors which would have led to defects. Perhaps the PO text template has fewer places for human errors. Lead time for the new design as well showed continuous improvement across runs.

The initial new design run presented incorrect pricing information and value, along with a minor title error in the contact information. The second run corrected the pricing value but retained the contact information title error. In the third run, while pricing referenced the correct document without specifying a value (and was therefore not considered a defect), the same title error persisted. The fourth run also had the title issue but no other defects. The fifth run saw a recurrence of title mistakes, with two errors present, although these were still considered minor. The recurring title problem indicates the prompt for Gemini, even after modifications, may still need adjustments to consistently pull information from the contract program.

Users must first extract required documents from zipped files, which can be time-consuming. These documents then need to be selected from your computer or Google Drive, and their names manually copied into the prompt. Furthermore, it is essential to review the generated text output after each run. This is because the text may show minor variations, and Google Gemini can occasionally produce information that is not accurately reflected in the source documents. However, despite of these challenges, the new design offers significant advantages. Overall quality is improved in most areas compared to the previous design. For example, the scope of work and the HSE requirements are explained more comprehensively.

This pilot case demonstrates the new design as a strong alternative to the old, offering a shorter lead time and more comprehensive text in several areas. The current prompt is highly usable, and minor refinements would likely establish the new design as the clearly superior choice.

RFQ Text

Table 4 presents a comparison of lead time and defects across five runs for both the old and new designs. The new design demonstrates a significant improvement in lead time, averaging 3 minutes and 15.6 seconds compared to the old design's average of 4 minutes and 54 seconds. This represents a reduction of 1 minute and 38.4 seconds, indicating that the new design is 33.47% more efficient in terms of lead time. Regarding defects, the new design averaged 0.4 defects per run, while the old design averaged 0.2 defects per run.

Table 4. Lead time and defects for the RFQ text.

	Lead time (minutes and seconds)		Defects (quantity)	
Run number	Old design	New design	Old design	New design
1	8 min 30 sec	4 min 5 sec	0	0
2	5 min 11 sec	2 min 37 sec	0	1
3	3 min 47 sec	2 min 42 sec	1	0

4	3 min 37 sec	3 min 35 sec	0	0
5	3 min 25 sec	3 min 19 sec	0	1

Lead time was reduced during the runs with the old design. Only one defect occurred, in the third run, due to my error with the contact information.

Lead time for new design decreased after the initial run but did not reduce further in subsequent runs. The first run was defect-free. A minor error, specifically an incorrect email address in the contract program document itself, was noted; this is considered a document error, not a defect by Gemini, and Gemini was unable to correct it. In the second run, Gemini omitted some names from the contact information, though the included names were accurate. Additionally, Gemini incorrectly generated one email address that was correct in the source document. The third run encountered no issues.

In the fourth run, a more comprehensive, and longer, text was generated. This time, Gemini inexplicably entered Canvas mode but successfully corrected an incorrect email address from the document. The fifth run did not trigger Canvas mode; Gemini still corrected the email address but exhibited a defect: it misinterpreted "Final date to confirm the bid" as the date a selected vendor accepts the selection, rather than the intended meaning of the deadline for vendors to confirm placing of the bid.

The new design offers improved comprehensiveness in many sections compared to the old. However, its output varies with each run, necessitating user caution. Defect analysis showed only one additional defect in the new design compared to the old, with no major defects present in the new version. While the 5-run average lead time improved with the new design, the lead times for recent runs were comparable to the old design. Overall, the new design presents a strong alternative to the old.

5 Conclusions

This study's concluding chapter provides an overview of the case company's current purchasing process, followed by an analysis of the pilot case outcomes, a discussion of managerial implications, a presentation of theoretical contributions to the field, an acknowledgment of the study's limitations, and suggestions for future research directions.

5.1 Summary of Findings

Feedback from the workshop sessions confirmed significant issues within the current process, including user dissatisfaction, problematic system integrations, inefficient data flow, and burdensome manual data entry tasks.

An evaluation determined that Google Gemini, in its current state, faces substantial limitations that hinder its ability to resolve these issues. Key constraints include lack of access to internal systems and data, difficulties processing large datasets or certain file types (like zipped files and folders), and insufficient contextual knowledge of the specific industry and business environment.

Furthermore, Gemini's current functionalities for tasks like automated presentation generation or precise template population from documents are not yet mature enough for reliable implementation. While the workshop didn't produce immediately actionable Gemini-based solutions due to these limitations, it successfully identified key process pain points and generated valuable concepts for future development, particularly as AI technology advances.

The choice of three pilot cases was suitable for the study's breadth. However, the similarity between the PO and RFQ text cases may have limited the generation of distinctly separate scientific insights.

The pilot testing phase highlighted several practical challenges in utilizing Google Gemini effectively for the intended tasks. A primary difficulty involved template population. Gemini required explicit, detailed instructions within the prompt and significant prompt refinement was necessary to receive the desired output. This raises a critical quality concern regarding the consistency and completeness of the generated output, as ensuring Gemini includes all information deemed essential by the company within the template proved problematic. Even when mostly successful, the final output structure often deviated from the precise template format. Furthermore, operational constraints were evident, as Gemini could not process entire folders of documents; input had to be curated manually, limited to a maximum of 10 individual files per request.

Despite these hurdles, the pilot results were mixed when evaluated. The pilot cases involving PO text and RFQ text generation demonstrated potential for significant lead time reduction. While they introduced only slightly more defects than the current process, a major weakness is the output inconsistency across different runs, making it difficult to fully trust the results without verification. Conversely, the pilot focused on generating an RFQ document list did not meet expectations. This case resulted in both increased lead times and a higher number of defects compared to the baseline.

The disparity in these results suggests Gemini may currently be better suited for more creative or generative tasks where adherence to strict formatting is less critical, rather than highly structured tasks like precise list generation from documents. The difficulty encountered with the RFQ document list, a task initially assumed to be straightforward for AI, underscores this observation about Gemini's current strengths and weaknesses.

This study managed to identify the current technical services purchasing process at the case company, waste and defects related to that process and potential ways how Google Gemini could mitigate that waste and defects. Potential pilot cases were identified, and the most prominent cases were selected. Based on the study, new design for pilot case RFQ document list fails to enhance the process step compared to the existing design, as lead time and defects are not reduced as an outcome. The results for PO and RFQ text were mixed as they did reduce the lead time, but defects were not reduced.

5.2 Contributions to Practice and Theory

This study provides valuable insights for the case company regarding the application of Google Gemini to its processes. While the research identified certain process issues that fall outside the scope of what Google Gemini or this study can directly fix, understanding these limitations is itself valuable for guiding future development efforts.

PO and RFQ text cases showcased promising results. It is recommended that these cases undergo further testing with end-users. If deemed successful, their implementation as formal process steps could be considered, although this would require acknowledging that the output quality and format differ from existing templates, necessitating changes in working methods and formal approval.

Consequently, the company now has clearer information regarding which tasks are suitable for Google Gemini and which are not, allowing for more focused research and development. Presenting these findings will also increase organizational awareness of Google Gemini's potential and limitations.

While it's wise not to expect large language models to resolve all process inefficiencies, they represent potentially valuable tools for the future. Therefore, staying informed about the evolution of this technology is crucial for the company to capitalize on readily achievable benefits as they arise.

This study provides a practical counterpoint to the often-optimistic portrayal of AI in procurement. Evaluating Google Gemini within a real-world purchasing process reveals that while such tools can offer efficiency gains, they remain susceptible to errors and limitations.

The challenges observed with Gemini, particularly its difficulty adhering strictly to processes and instructions compared to its capacity for unguided creativity, may be indicative of broader issues with current LLMs. This suggests similar limitations could arise when implementing LLMs in other process-reliant fields like procurement that demand rigorous compliance with established procedures.

This realism prompts a necessary evaluation: Are the current benefits derived from LLMs sufficient to justify the significant investment and the associated social and environmental concerns, especially given their practical shortcomings in structured environments?

5.3 Limitations and Further Research

The evolving nature of Gemini, with its potential for behavioral changes over time, presents a challenge for long-term studies. The workshop sessions provided substantial background information on process operations but offered limited direct input for case identification, as potential cases were largely identified during the preceding process definition phase. The subsequent case selection, which utilized a Pugh Matrix based on the purchasing process analysis, lacked fully detailed scoring criteria and explicit justifications for each score assigned. Furthermore, the measurement phase was constrained by time, permitting only five runs per design. A larger sample size would have enhanced the reliability of the sample-based estimates.

Investigating how Google Gemini features such as NotebookLM, Canvas, and Deep Research can be effectively applied is valuable. While the current pilot cases demonstrate potential, their outcomes can be improved with better prompts. It's also important to recognize that other potential cases, not selected as pilots, may hold untapped value for process improvement.

The ongoing, rapid development of LLMs like Google Gemini is a key factor. As newer versions become available, they will likely introduce enhanced capabilities relevant to process optimization, including purchasing. The extent to which Gemini can improve future processes hinges on advancements in its ability to follow strict rules, access systems, and manage large-scale data and documents effectively.

References

Aljazzazen, S., & Schmuck, R. (2022). *Critical success factors for successful Lean Six Sigma implementation in the service organizations.* Quality - Access to Success, 23(188), 76-85. https://doi.org/10.47750/QAS/23.188.11

Allal-Cherif, O., Simon-Moya, V., & Cuenca Ballester, A. C. (2021). Intelligent purchasing: How artificial intelligence can redefine the purchasing function. Journal of Business Research, 124, 69-76. https://doi.org/10.1016/j.jbusres.2020.11.050

Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M. A., Al-Amidie, M., & Farhan, L. (2021). Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. *Journal of Big Data, 8*(1), 53. https://doi.org/10.1186/s40537-021-00444-8

Antony, J. (2006). Six Sigma for service processes. *Business Process Management Journal, 12(2),* 234-248. https://doi.org/10.1108/14637150610657558

Antony, J. (2016). *Lean Six Sigma for Small and Medium Sized Enterprises: A Practical Guide.* CRC Press.

Antony, J., Snee, R., & Hoerl, R. (2017). Lean Six Sigma: Yesterday, today and tomorrow. *International Journal of Quality & Reliability Management*, *34*(7), 1073–1093. https://doi.org/10.1108/IJQRM-03-2016-0035

Axelsson, B., & Wynstra, F. (2002). *Buying business services*. John Wiley.

Basu, R. (2011). *Fit Sigma: A Lean Approach to Building Sustainable Quality Beyond Six Sigma.* Wiley.

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency,* 610-623. https://doi.org/10.1145/3442188.3445922

Benton, W. C. (2020). *Purchasing and supply chain management* (4th ed.). SAGE Publications, Incorporated.

Bertalanffy, L. von. (1968). *General system theory: Foundations, development, applications.* Braziller.

Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S.,... & Liang, P. (2022). On the opportunities and risks of foundation models. https://doi.org/10.48550/arXiv.2108.07258

Bostrom, N. (2014). *Superintelligence: Paths, dangers, strategies.* Oxford University Press.

Breyfogle, F. W. (2003). *Implementing Six Sigma: Smarter Solutions Using Statistical Methods* (2th ed.). Wiley.

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., … Amodei, D. (2020). Language models are few-shot learners. OpenAI. https://doi.org/10.48550/arXiv.2005.14165

Brynjolfsson, E., & McAfee, A. (2017). *Machine, platform, crowd: Harnessing our digital future* (First edition.). W.W. Norton & Company.

Burger, M., Nitsche, A., & Arlinghaus, J. (2023). Hybrid intelligence in procurement: Disillusionment with AI's superiority? *Computers in Industry, 150*, 103946. https://doi.org/10.1016/j.compind.2023.103946

Charles, V., Emrouznejad, A., & Gherman, T. (2023). A critical analysis of the integration of blockchain and artificial intelligence for supply chain. *Annals of Operations Research, 327*(1), 7–47. https://doi.org/10.1007/s10479-023-05169-w

Christiano, P. F., Leike, J., Brown, T., Martic, M., Legg, S., & Amodei, D. (2017). Deep reinforcement learning from human preferences. *Advances in Neural Information Processing Systems, 30,* 4765-4773. https://doi.org/10.48550/arXiv.1706.03741

Christopher, M. (2016). *Logistics & supply chain management* (5th ed.). Pearson Education Limited.

Culot, G., Podrecca, M., & Nassimbeni, G. (2024). Artificial intelligence in supply chain management: A systematic literature review of empirical studies and research directions. Computers in Industry, 162, 104132. https://doi.org/10.1016/j.compind.2024.104132

Davenport, T. H., & Short, J. E. (1990). The new industrial engineering: Information technology and business process redesign. *Sloan Management Review, 31(4),* 11-27.

Deepali Kishor Desai. (2010). *Six sigma*. Himalaya Publishing House.

Deming, W. E. (2000). *Out of the crisis.* MIT Press.

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). *BERT: Pre-training of deep bidirectional transformers for language understanding. arXiv.* https://doi.org/10.48550/arXiv.1810.04805

Feigenbaum, E. A. (1984). Knowledge engineering: The applied side of artificial intelligence. *Annals of the New York Academy of Sciences, 426(1),* 91-107.

Gemini Team. (2023). *Gemini: A family of highly capable multimodal models*. Google.

Google. (2023, December 6). Introducing Gemini: Google's most capable AI model yet. Google Blog. Retrieved 26.1.2025: https://blog.google/technology/ai/google-gemini-ai/

Guida, M., Caniato, F., Moretto, A., & Ronchi, S. (2023). The role of artificial intelligence in the procurement process: State of the art and research agenda. *Journal of Purchasing & Supply Management, 29*, 100823. https://doi.org/10.1016/j.pursup.2023.100823

Hammer, M., & Champy, J. (1993). Reengineering the corporation: A manifesto for business revolution. *Business Horizons, 36*(5), 90–91. https://doi.org/10.1016/S0007-6813(05)80064-3

Heinis, S., Bamford, D., Papalexi, M., & Vafadarnikjoo, A. (2022). Services procurement: A systematic literature review of practices and challenges. *International Journal of Management Reviews, 24(3),* 352-372. https://doi.org/10.1111/ijmr.12281

Hofmann, E., Brunner, J. H., & Holschbach, E. (2020). Research in business service purchasing: Current status and directions for the future. *Management review quarterly, 70*(3), 421-460. https://doi.org/10.1007/s11301-019-00172-7

https://doi.org/10.48550/arXiv.2312.11805

Kraljic, P. (1983, September). Purchasing must become supply management. *Harvard Business Review.* Retrieved 4.2.2025 from https://hbr.org/1983/09/purchasing-must-become-supply-management

Le, W. T., Maleki, F., Romero, F. P., & Forghani, R. (2020). Overview of machine learning part 1: Fundamentals and classic approaches. *Neuroimaging Clinics of North America, 30*(4), 433–445. https://doi.org/10.1016/j.nic.2020.08.004

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature (London), 521*(7553), 436-444. https://doi.org/10.1038/nature14539

Lee, H.-P., Sarkar, A., Tankelevitch, L., Drosos, I., Rintel, S., Banks, R., & Wilson, N. (2025). *The impact of generative AI on critical thinking: Self-reported reductions in cognitive effort and confidence effects from a survey of knowledge workers*. In *CHI Conference on Human Factors in Computing Systems (CHI '25), April 26–May 01, 2025, Yokohama, Japan*. ACM. https://doi.org/10.1145/3706598.3713778

Liker, J. K. (2004). *The Toyota Way: 14 Management Principles from the World's Greatest Manufacturer.* McGraw-Hill.

Lysons, K., & Farrington, B. (2020). *Procurement and supply chain management* (10th ed.). Pearson Education Limited.

McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (2006). A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence: August 31, 1955. *The AI magazine, 27*(4), 12-14. https://doi.org/10.1609/aimag.v27i4.1904

Newell, A., & Simon, H. A. (1956). The logic theory machine–A complex information processing system. *IRE Transactions on Information Theory, 2(3),* 61-79. https://doi.org/10.1109/TIT.1956.1056797

Nonthaleerak, P., & Hendry, L. (2008). Exploring the six sigma phenomenon using multiple case study evidence. *International Journal of Operations & Production Management, 28*(3), 279-303. https://doi.org/10.1108/01443570810856198

Ohno, T. (1988). *Toyota production system: Beyond large-scale production.* Productivity Press.

Panchal, R. K., Bhoite, S. B., Pande, A., Patil, R., & Mukkemwar, R. (2024, November). *Comprehensive study of Google Gemini and text-generating models: Understanding capabilities and performance.* Conference Paper presented at Grenze International Journal of Engineering and Technology. Retrieved from 12.2.2025 https://www.researchgate.net/publication/386101230

Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A Design Science Research Methodology for Information Systems Research. *Journal of management information systems, 24*(3), 45-77. https://doi.org/10.2753/MIS0742-1222240302

Poulain, H. (Director). (2024). *In the belly of AI* [Film]. Federation Studios.

Pugh, S. (1991). *Total Design: Integrated Methods for Successful Product Engineering*. Addison-Wesley Publishing Company.

Pyzdek, T., & Keller, P. A. (2014). *The Six Sigma handbook* (4th ed.). McGraw-Hill Education.

Quarshie, A. M., Salmi, A., & Leuschner, R. (2016). Sustainability and corporate social responsibility in supply chains: The state of research in supply chain management and business ethics journals. *Journal of Purchasing and Supply Management, 22*(2), 82–97. https://doi.org/10.1016/j.pursup.2015.11.001

Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). *Language models are unsupervised multitask learners.* OpenAI. https://cdn.openai.com/better-language-models/language_models_are_unsupervised_multitask_learners.pdf

Rosenbaum, M. S., Otalora, M. L., & Ramírez, G. C. (2017-01). How to create a realistic customer journey map. *Business horizons, 60*(1), 143-150. https://doi.org/10.1016/j.bushor.2016.09.010

Russell, S. J. & Norvig, P. (2022). *Artificial intelligence: A modern approach* (Fourth edition. Global edition.). Pearson.

Samanta, M., Virmani, N., Singh, R. K., Haque, S. N., & Jamshed, M. (2024). Analysis of critical success factors for successful integration of lean six sigma and Industry 4.0 for organizational excellence. *TQM journal, 36*(1), 208-243. https://doi.org/10.1108/TQM-07-2022-0215

Sennrich, R., Haddow, B., & Birch, A. (2016). Neural machine translation of rare words with subword units. *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (ACL 2016)*, 1715–1725. 10.18653/v1/P16-1162

Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., van den Driessche, G., . . . Hassabis, D. (2016). Mastering the game of Go with deep neural networks and tree search. *Nature (London), 529*(7587), 484. https://doi.org/10.1038/nature16961

Spreitzenbarth, J. M., Bode, C., & Stuckenschmidt, H. (2024). Artificial intelligence and machine learning in purchasing and supply management: A mixed-methods review. *Journal of Purchasing & Supply Management, 30*, 100896. https://doi.org/10.1016/j.pursup.2024.100896

Taylor, F. W. (1911). *The principles of scientific management.* Harper & Brothers.

Toorajipour, R., Sohrabpour, V., Nazarpour, A., Oghazi, P., & Fischl, M. (2021). Artificial intelligence in supply chain management: A systematic literature review. Journal of Business Research, 122, 502-517. https://doi.org/10.1016/j.jbusres.2020.09.009

Turing, A. M. (1950). Computing machinery and intelligence. *Mind, 59(236),* 433-460.

van der Aalst, W. (2012). Process Mining: Overview and Opportunities. *ACM transactions on management information systems, 3*(2), 1-17. https://doi.org/10.1145/2229156.2229157

Van der Valk, W., & Rozemeijer, F. (2009). Buying business services: Towards a structured service purchasing process. *The Journal of Services Marketing, 23*(1), 3–10. https://doi.org/10.1108/08876040910933048

Wamba, S. F., Guthrie, C., Queiroz, M. M., & Minner, S. (2024). ChatGPT and generative artificial intelligence: An exploratory study of key benefits and challenges in operations and supply chain management. International Journal of Production Research, 62(16), 5676-5696. https://doi.org/10.1080/00207543.2023.2294116

Womack, J. P., & Jones, D. T. (1997). Lean Thinking—Banish Waste and Create Wealth in your Corporation. *Journal of the Operational Research Society*, *48*(11), 1148. https://doi.org/10.1057/palgrave.jors.2600967

Zellers, R., Holtzman, A., Clark, C., Lu, D., & Ramasamy, K. (2019). Defending Against Neural Fake News. *Proceedings of the 2019 Conference on Neural Information Processing Systems*, 2291–2301. https://doi.org/10.48550/arXiv.1905.12616

Appendices

Appendix 1. Development comments from the workshop

Problem Area	Details
All	Multiple systems (Project management system, tendering system, ERP system) with their own approval processes, slowing down operations and causing duplicate work. Challenges in transferring data between systems.
All	System usage is challenging for worksite managers. Design team packages and instructions are inadequate.
All	Designer and CM (Construction Managers) don't consistently upload documents to Document Management system.
All	Lack of clear job descriptions or responsibility matrices.
All	HSE guideline version updates are problematic (too many notifications, specs not readily available).
All	Contractors may not be familiar with frame agreements.
Bidding process	Variable quality of TMRs (schedule, documents, tendering). Necessary documents may not be accessible via the provided link. RFQ documents may be scattered across different locations.
Bidding process	TMR creation has shifted from the worksite to the design team, causing documentation issues. Obtaining complete TMRs is rare. If a TMR is missing, orders are placed in other systems (e.g., ERP system).
Bidding process	Lack of or unclear procurement plan. Missing or inadequate scoring criteria creation.
Bidding process	Issues sending RFQs from tendering system due to large file sizes. Security concerns with using Google Drive. Slow manual processing of RFQ letters.

Bidding process	Challenges in ensuring comparability of bids. Large number of evaluations complicating comparison. Suppliers not always providing transparent information. Difficulty comparing bids.
Bidding process	Stakeholders not actively participating in scoring. Lack of historical data on additional work.
Bidding process	Tendering system cannot handle large document uploads. Request for quotation material size exceeds tendering system's limitations.
Bidding process	Excessive repetition in negotiation notes. Questionable quality of negotiation templates.
Bidding process	Lack of numerical values in scoring forms.
Bidding process	Price comparison forms (especially for smaller contracts) are different, and tendering system doesn't provide meaningful price comparisons.
Bidding process	Tendering system's questionnaire responses are not sufficiently detailed or reliable (e.g., statements about staffing levels).
Bidding process	Negotiations update information, causing tendering system data to become outdated.
Bidding process	Contractor selection, price comparison, and negotiations lack consistency.
Bidding process	TMR (Technical Material Requisition) doesn't arrive in Project management system on time or is incomplete/incorrect.
Bidding process	Delays in TMR or approval notifications. Information may arrive too late. Poor communication between the worksite and the design team.
Managing purchase orders	Work start initiation requires written pre-approval and a 30-day waiting period from the order date.

Managing purchase orders	Work schedule and Project management system start dates don't match.
Managing purchase orders	PO attachment management is challenging (especially changes).
Managing purchase orders	Lack of contract terms presentation to supervisors before work begins, leading to later issues. Contract terms updates (e.g., schedules, work specifications) require full contract revision.
Managing purchase orders	Subcontractor liability document verification should be done before the order.
Managing purchase orders	Subcontractor auditing is inadequate and unsupervised (e.g., contractor statements).
Managing purchase orders	The Purchaser is merely an approver of subcontractor approvals, even though procurement may have information on prohibited subcontractors. The Purchaser might not remember prohibited subcontractors. Why is it even possible to approve prohibited subcontractors
Managing purchase orders	Contractor Management system performance is slow.
Managing purchase orders	Security deposit management in invoice verification is inefficient and unclear.
Managing purchase orders	AI use for invoice verification (e.g., unit prices, Report of the Work data) could streamline the process.

Managing purchase orders	Invoices often have errors, and credit notes are delayed.
Managing purchase orders	Work site meetings record additional and change works, but Project management system may not be updated.
Managing purchase orders	Additional work specifications should provide more detailed information (including price lists).
Purchase order closing	Non-stock delivery function doesn’t always work correctly in Project management system.
Purchase order closing	ERP system PO's may not close automatically.
Purchase order creation	Manual PO processing is time-consuming (PO text, payment installments, attachments).
Purchase order creation	PO approval process is slow and may contain errors (approvers, missing personnel). POs under 50,000 euros should be sent automatically to vendor but this is not possible in the ERP system currently.