Engineering Reports

WILEY

Engineering Reports Open Access

REVIEW OPEN ACCESS

Large Language Model-Based Cognitive Assistants for Quality Management Systems in Manufacturing: A Requirement Analysis

Marcos Galdino | Tobias Hamann | Anas Abdelrazeq | Ingrid Isenhardt

IQS Intelligence in Quality Sensing, Laboratory for Machine Tools and Production Engineering (WZL) of RWTH Aachen University, RWTH Aachen University, Aachen, Germany

Correspondence: Marcos Galdino (marcos.galdino@wzl-iqs.rwth-aachen.de)

Received: 20 May 2025 | **Revised:** 24 July 2025 | **Accepted:** 28 September 2025

Funding: This work was supported by German Federal Ministry of Labour and Social Affairs (BMAS) and the European Union via the European Social Fund Plus (ESF Plus) as part of the "Future Centres" program within the project ZUKIPRO [ZUK-1-0005].

Keywords: cognitive assistants | digital transformation | human-technology-organization | ISO 9001 | large language models | manufacturing | quality management systems

ABSTRACT

The integration of large language model-based cognitive assistants (LLM-CAs) into manufacturing offers opportunities to enhance decision-making, quality, and efficiency. However, aligning LLM-CAs with quality management systems (QMS), such as ISO 9001:2015, remains a complex task. This study systematically reviewed 53 studies (2022–2024) to identify 84 literature-based requirements related to LLM-CAs deployment for QMS. We translated the identified challenges into 24 actionable requirements using thematic analysis. Grounded in the human, technology, and organization concept (HTO), we classified the requirements across HTO subsystems and their overlaps and mapped them to the seven ISO 9001:2015 clauses. Our analysis reveals strong alignment between LLM-CAs and QMS principles, particularly in areas such as support in decision-making processes and continuous improvement. Moreover, our findings highlight persistent challenges, such as transparency, compliance risks, and workforce adaptation. An illustrative case study of setup time optimization demonstrates the practical application of these findings. The results provide a structured foundation for QMS-compliant integration of LLM-CAs into manufacturing. Future research should extend these findings through stakeholder-driven validation, system architecture development, and real-world implementation studies. This extension would, therefore, support responsible and human-centered LLM-CA adoption in digitalized manufacturing environments.

1 | Introduction

The integration of large language model-based cognitive assistants (LLM-CAs) into manufacturing presents a significant opportunity to enhance operational performance. This integration is particularly relevant when aligned with established quality management systems (QMS), such as ISO 9001:2015.

Large language models (LLMs) are algorithms pre-trained on extensive, heterogeneous data corpora. Thereafter, they are fine-tuned with specific data to optimize context alignment and performance. Drawing from diverse data sources, for example, web texts, academic literature, and code, LLMs acquire knowledge and demonstrate capabilities in advanced data processing and interpretation [1].

Several studies illustrate the application potential of LLMs. Jin et al. [2] highlight their capability in time series forecasting. Paroha and Chotrani [3] emphasize their potential in predictive maintenance to enhance operational efficiency and reduce unplanned downtimes. Additionally, Kernan Freire et al. [4] demonstrate how LLM-CAs can collaborate with shop floor employees, mimicking human-like cognitive functions, for example, communication and interaction with multiple users.

These LLM-CA applications are especially relevant in manufacturing, a field undergoing rapid digital transformation driven by technologies such as artificial intelligence (AI) [5]. Manufacturing involves complex systems that integrate production processes, human resources, machinery, and technology [6]. In this context, digitalization transforms individual work tasks and enables new business models, while digitalizing processes and driving innovative solutions that enhance customer value [5]. Additionally, digitalization can lead to improvements in process efficiency through information and communication technologies [7].

QMS provides the foundational structure for manufacturing operations. As defined by the International Organization for Standardization (ISO), its QMS framework supports consistent product quality, regulatory compliance, and continuous improvement [8]. Adopting QMS can bring benefits like: "(a) the ability to consistently provide products and services that meet customer and applicable statutory and regulatory requirements; (b) facilitating opportunities to enhance customer satisfaction; (c) addressing risks and opportunities associated with its context and objectives; (d) the ability to demonstrate conformity to specified quality management system requirements" [8, p. 6]. This approach drives operational improvements, for example, cost reductions and increased work efficiency [9].

Despite the potential synergy between LLM-CAs, QMS, and manufacturing, several challenges hinder integration. These challenges include lack of data privacy and security [10–14], lack of skilled workforce [15–17], hallucination [18–20], among others. Identifying these challenges and translating them into actionable system requirements is essential for enabling the deployment of LLM-CAs for QMS in manufacturing.

Although potentially beneficial, this integration leads to increased work systems complexity. The human, technology, and organization (HTO) concept offers a suitable approach for this purpose. It emphasizes the interrelation of human, technological, and organizational subsystems in shaping effective work environments [21, 22].

Although relevant, current research lacks a structured classification of LLM-CA requirements based on the HTO concept, particularly for QMS in manufacturing. On that basis, a multisystem approach is needed to assess the feasibility of this integration. As an example, Ghimire et al. [14] stress the importance of aspects such as workforce training, copyright and intellectual property, and the collaborative involvement of industry experts and policymakers. Similarly, Doanh et al. [23] call for interdisciplinary studies that integrate researchers and practitioners from varied disciplines, that is, technical, social, economic, and ethical. Both authors identify these needs as directions for future research, but do not address them in their current work.

Bridging this gap is crucial for enabling the deployment of LLM-CAs for QMS in manufacturing. This study addresses the gap by answering the following main research question (MRQ) and its sub-research questions (SRQ):

MRQ. *How can HTO-aligned requirements support the deployment of LLM-CAs for QMS in manufacturing?*

SRQ 1. *What are the key challenges in deploying LLM-CAs for QMS in manufacturing?*

SRQ 2. *How can the identified challenges be translated into requirements according to the HTO concept?*

SRQ 3. *How can the interaction of the derived requirements leverage LLM-CAs for QMS in manufacturing?*

The goal of this study is to provide a holistic and structured classification of literature-based requirements to support the integration of LLM-CAs for QMS in manufacturing grounded in the HTO concept (Figure 1). To achieve this goal, we first identify key challenges related to LLM-CAs for QMS. Next, we translate the identified challenges into HTO-aligned requirements. Finally, we analyze how the derived requirements interact to leverage QMS under ISO 9001:2015.

Throughout this article, we use the acronym *LLM-CAs for QMS* to refer to "large language model-based cognitive assistants for quality management systems in manufacturing."

This study targets researchers, industry practitioners, and policymakers working at the intersection of artificial intelligence, QMS, and manufacturing. It is particularly relevant to professionals and researchers involved in digital transformation and operational excellence. By providing a structured classification of requirements for integrating LLM-CAs for QMS, this study offers a baseline for developing human-centered, technology-driven solutions.

This paper is structured as follows: Section 2 presents the theoretical background, covering key concepts and theories related to LLM-CAs, manufacturing, QMS, and HTO. Section 3 describes the methods. Section 3.1 details the systematic literature review (SLR) used to find relevant studies and challenges. Section 3.2

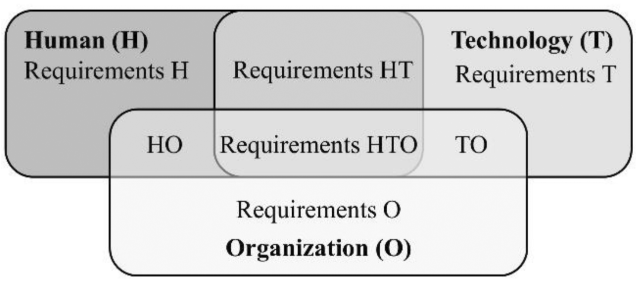

FIGURE 1 | Schematic classification of requirements according to the HTO concept.

outlines the Thematic Analysis (TA), explaining how the challenges identified were turned into requirements and then associated with the HTO concept. Section 4 presents the results, providing a detailed classification of the identified requirements. The discussion in Section 5 interprets these results in the context of ISO 9001:2015 and showcases application through a case study. Finally, Section 6 concludes the paper with a summary of the findings and an outlook on future research directions and potential practical and theoretical applications.

2 | Theoretical Background

2.1 | Large Language Model-Based Cognitive Assistants

LLMs revolutionized natural language processing (NLP) by using advanced architectures and vast datasets to enable machines to process and generate human-like text [24]. A key breakthrough came in 2017 when Vaswani et al. [25] introduced the Transformer architecture, now the foundation of most LLMs. Unlike earlier models that processed text sequentially, the Transformer uses attention mechanisms to capture dependencies across entire texts. This design supports parallel processing, making it easier to scale and train models on extensive data corpora [25].

One of the most prominent Transformer-based models is GPT-4, developed by OpenAI [26]. GPT-4 demonstrates exceptional performance on human-designed tests, for example, it achieved a top 10% score on a simulated bar exam (licensing test for aspiring lawyers), surpassing most human test-takers.

LLMs are pre-trained on diverse datasets, for example, websites, academic texts, and code, which enables them to generalize across tasks and domains. This flexibility is evident in their ability to perform zero-shot, one-shot, and few-shot learning [27]. In zero-shot learning, LLMs complete tasks without any previous examples. In one-shot learning, one example supports model task completion, while few-shot learning involves multiple examples. For instance, Li et al. [28] proposed a few-shot learning approach for code-to-code translation, that is, from Java to Python and from Java to C#. In contrast, fine-tuning updates the weights of pre-trained models to particular tasks using labeled data [27]. Chen et al. [1], for example, developed Data-Juicer, which is a system that enables both pre-training and fine-tuning by generating custom data recipes.

These previous examples reflect human-like capabilities of LLMs, particularly when used as cognitive assistants in industrial settings. Kernan Freire [29] emphasizes that LLMs can support knowledge acquisition, sharing, and application through conversational human–machine interactions. Figliè et al. [30] proposed an LLM industrial cognitive assistant for human-centered applications. The authors designed an LLM-CA to support managers and operators in their daily activities through the connection of a chatbot and machines. Initial user feedback confirmed the assistant's potential but also revealed limitations such as uncertainty about data sources and the credibility of the content generated.

Despite their disruptive aspect, LLM-CAs still face several challenges. These include difficulties in integrating with legacy and heterogeneous factory systems [7], reasoning errors [4], and limited real-time responsiveness [7, 12]. Furthermore, deployment in manufacturing environments requires attention to human and organizational factors, besides technological ones. For instance, researchers identified the need to address artificial intelligence avoidance [16, 31, 32] and the lack of a skilled workforce [15–17].

In conclusion, LLMs, such as GPT-4 by OpenAI [26] have transformed NLP through the Transformer architecture. In manufacturing, LLM-CAs can enhance worker competencies by supporting knowledge acquisition and application. However, their effective integration depends on addressing human, technological, and organizational challenges. These are best understood through the lens of the HTO concept.

2.2 | Manufacturing

Manufacturing is a complex system that integrates production processes and systems, human labor, machinery, and technology [5]. As Jones et al. [33] describe, the manufacturing industry functions like a dynamic, evolving organism that innovates and produces, not being a rigid and mechanized system. Moreover, it plays a vital role in national economies by contributing significantly to GDP, employment, innovation, science and technology, and education [34].

Shabur [35] emphasizes that, since 2011, when the German government coined the term Industry 4.0, manufacturing has been undergoing a digital revolution. Industry 4.0 refers to the digitalization of production processes and entire value chains. It interconnects people, processes, and objects through digital systems via real-time data exchange via the internet [36]. Suleiman et al. [37] argue that by integrating key technologies, for example, AI, Industry 4.0 disrupts both production processes and manufacturing systems. According to the World Economic Forum [38], the adoption of AI, cloud computing, and big data remains a top priority throughout industries. Seventy-five percent of companies plan to adopt them by 2028 [38].

The impact of Industry 4.0 is already visible in practice. For example, a Mercedes-Benz plant in Brazil integrated Azure Machine Learning into its sales operations to generate more accurate, data-driven sales recommendations. The solution combined internal data, for example, sales records, with external data, for example, macroeconomic indicators, to tailor its sales offers [39].

As Fonseca et al. [40] note, the key features of Industry 4.0 are digitalization, automation, customization, flexibility, short time-to-market, human–machine interaction, automatic data exchange and communication, and decentralized decision-making. However, the transformation goes beyond technical upgrades. Calış Duman and Akdemir [41] argue that Industry 4.0 also requires a fundamental rethinking of organizational structures, business models, and work methods. The reorientation of business models and corporate strategies will shift work tasks and dissolve strict structural limits in companies. Therefore, Industry 4.0 fosters flexible organizational structures because of the increased permeability of business areas [41].

At the employee level, Industry 4.0 introduces novel forms of work. Fettig et al. [42] highlight that Industry 4.0 enables self-dependent work, decision and location independence, and time-flexible working arrangements. Fantini et al. [43] emphasize the emergence of a symbiotic relationship between employees and technological systems. In this context, employees contribute with human capabilities hardly replaceable by artificial intelligence and automation. These capabilities encompass, for instance, data-driven decision-making, conceptualization, and idealization of innovative processes and projects [43].

Yet, there are also workforce challenges. As Sony [44] states, Industry 4.0 demands a highly skilled workforce with technical, e.g., coding, methodological, e.g., problem-solving, social, e.g., teamwork, and personal, e.g., adaptability, competencies. Meeting these demands requires targeted training programs and strategic investments in workforce development, which may lead to increased organizational costs [44].

In summary, Industry 4.0 is reshaping manufacturing through digital technologies, leading to greater efficiency, flexibility, and innovation. At the same time, it requires a reconfiguration of organizational structures and a highly skilled workforce. Bridging these demands requires strategic training programs coupled with strategic technology implementation.

2.3 | Quality Management Systems as Per ISO 9001:2015

Quality management systems, as defined by ISO 9001:2015, provide a structured framework to ensure consistent quality, regulatory compliance, and organizational excellence [8]. According to Hoyle et al. [45], QMS help organizations showcase their ability to meet customer needs while committing to continuous improvement. Originally released in 1987 by the International Organization for Standardization (ISO), ISO 9001 aimed at harmonizing fragmented quality standards, especially within manufacturing, and at supporting global trade [46]. The version ISO 9001:2015 emphasizes risk-based thinking, leadership commitment, stakeholder engagement, process approach, change and knowledge management, and continuous improvement [47].

ISO 9001:2015 [8, p. 2] defines QMS as follows:

1. A QMS comprises activities by which the organization identifies its objectives and determines the processes and resources required to achieve desired results.
2. The QMS manages the interacting processes and resources required to provide value and realize results for relevant interested parties.
3. The QMS enables top management to optimize the use of resources considering the long- and short-term consequences of their decision.
4. A QMS provides the means to identify actions to address intended and unintended consequences in providing products and services.

ISO 9001:2015 comprises seven quality management principles: customer focus, leadership, engagement of people, process approach, improvement, evidence-based decision-making, and relationship management [8]. These principles are operationalized through seven clauses, as per Figure 2. The clauses offer a step-by-step structure for designing, implementing, and improving a QMS.

The seven clauses ensure that QMS strategies align with an organization's strategic direction and customer needs [48]. Empirical evidence supports the value of QMS. For instance, Solomon et al. [49], through quantitative research, found that the dissemination of QMS improves the electricity industry. According to them, QMS positively affected attitudes towards securing a sustainable long-term energy supply. It improved processes, fostered a quality culture, and contributed to continuous improvement.

Beyond traditional industries, QMS are also being adapted for emerging technologies. Mustroph and Rinderle-Ma [50] proposed a QMS concept delivered as Software as a Service (SaaS), aligned with the European Union Artificial Intelligence Act (EU AI Act). The software's goal was to monitor, document the design, and manage AI system quality and risk. Their proof-of-concept showed the incorporation of data from a sub-service into technical documentation. This example highlights the growing applicability of QMS to domains like AI, extrapolating traditional QMS application fields.

Despite their strengths, QMS implementation is not without challenges. Fadilasari et al. [51] classify these challenges into three categories:

1. Ecosystem enablers, for example, innovation ecosystem and people involvement;
2. Information technology and governance, for example, cyber security and data protection; and

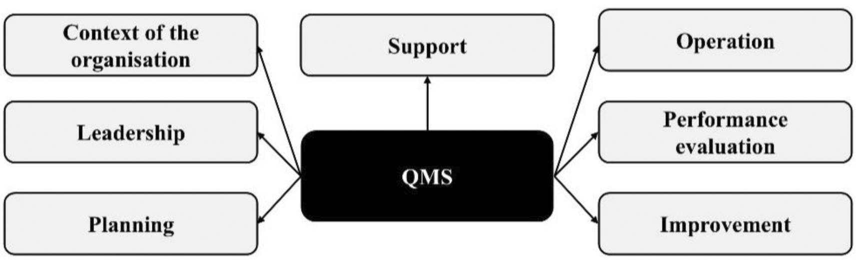

FIGURE 2 | Clauses of the ISO 9001:2015 quality management systems—requirements.

3. Organizational behavior, for example, resistance to change and leadership.

Nevertheless, the adoption of QMS is vital to foster the competitiveness of companies, especially in Industry 4.0, where digitalization and quality are interdependent [52].

In summary, ISO 9001:2015 offers a comprehensive framework for ensuring product and service quality while promoting adaptability and continuous improvement. As industries evolve, particularly under Industry 4.0, the role of QMS becomes even more vital, bridging the gap between operational excellence and the demands of modern digital systems.

2.4 | Human, Technology, Organization Concept (HTO)

The HTO concept provides a structured approach for understanding, analyzing, and designing human work systems. As defined by Berglund et al. [21], HTO emphasizes the interplay among three subsystems: human factors, e.g., tacit competences, technological tools, for example, software, and organizational structures, for example, management style [21]. Rooted in sociotechnical systems theory, which emerged in the mid-20th century, the concept views these three subsystems as interconnected [22]. On that basis, they shape the system's performance, health issues, and safety [53, 54].

Karltun et al. [22] state that each subsystem, that is, human, technology, and organization, has its own set of requirements. Yet, these subsystems interact inseparably in real-world work activities. For instance, effective task execution depends not only on the capabilities of human operators but also on supporting technologies and the surrounding organizational context [22]. By accounting for these interactions, HTO facilitates the design of work systems that are not only efficient but also supportive of human needs [21].

In industrial settings, the HTO concept applies to guiding both technological deployment and organizational changes. Aas and Johnsen [55], for example, used the HTO concept to address human factors in the design of control centers in petroleum projects. They argued that ISO 11064 (Ergonomic design of control centers) should better incorporate the HTO elements. These include organizational roles and competence development to improve usability and performance.

Similarly, Hallén et al. [56] applied the HTO concept in Building Information Modeling (BIM). They used BIM as a process of creating and managing information for a building asset. Their results revealed that successful BIM implementation requires more than just robust technology, as the acceptance of BIM plays an essential role in its deployment. They concluded that "BIM is a holistic and social system" [56, p. 1], which is better captured through the HTO lens.

In summary, the HTO concept offers a valuable human-centered perspective for understanding, analyzing, and designing work systems [21]. By accounting for the interdependence of human, technological, and organizational subsystems, HTO helps ensure that new developments, for example, LLM-CAs, are both effective and accepted. The concept's application across industries illustrates its potential for guiding integrated system design in rapidly evolving technological contexts [55, 56].

3 | Methods

This study employs two distinct but complementary methods to achieve its goals. The first of them is the systematic literature review according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) [57]. The second method is the Thematic Analysis [58], which aims at analyzing qualitative data. We detail the former in Section 3.1 and the latter in Section 3.2.

3.1 | SLR—PRISMA 2020

Following Moher et al. [59], this study adopted the PRISMA 2020 method to guide the SLR. PRISMA offers a structured approach for conducting and reporting reviews and meta-analyses of other types of research, for example, qualitative or quantitative studies. PRISMA 2020 comprises steps for identifying, screening, and synthesizing literature evidence [59], consisting of a 27-item checklist and a four-phase flow diagram.

This study does not detail each of the 27 items. Instead, it focuses on the most relevant ones for achieving the research objectives. A comprehensive explanation of the full checklist is present in the cited references [57, 59–61]. In PRISMA 2020, "records" refer to any source identified during the search phase, for example, abstracts, gray literature, while "studies" refer to actual research described in the records, for example, peer-reviewed articles.

The following four phases underpin the method [59]:

1. Identification: identification of records through database searching.
2. Screening: selection of records after removal of duplicates and screening of title or abstract for relevance.
3. Eligibility: selection of full-text articles for analysis after exclusion of irrelevant ones.
4. Included: number of studies included in the analysis.

PRISMA 2020 helps standardize the review process, ensuring that inclusion criteria, data extraction, and synthesis are conducted transparently and rigorously [61].

3.1.1 | Eligibility Criteria

To guide the selection of studies, we derived the following keywords, as per Table 1, stemming from the most relevant terms in the MRQ, that is, LLMs, cognitive assistants, and manufacturing. Given its close association with industrial contexts, we included the term QMS as a parent keyword for manufacturing.

Drawing from the keywords, we constructed the following Boolean search string (the special character "*" serves to include variations of the same word, e.g., model* = model & models).

("Large Language Model*" **OR** "LLM" **OR** "Conversation* AI Model*" **OR** "Generative Pre-trained Transformer Model*" **OR** "GPT" **OR** "Generative AI" **OR** "GenAI" **OR** "Generative Artificial Intelligence Model*") **AND** ("Cognit* Assistant*" **OR** "Conversation* AI" **OR** "AI Assistant*" **OR** "Machine Learning Agent*" **OR** "Assistant*" **OR** "Chatbot*") **AND** ("Manufactur*" **OR** "Quality Management System*" **OR** "Industry 4.0" **OR** "Fabric*" **OR** "Process*" **OR** "Product*").

Table 2 identifies all the inclusion and exclusion criteria for the selection of studies for synthesis. These criteria encompass, together with the search string, the eligibility criteria.

3.1.2 | Information Sources

We chose four information sources for the searching of records, as per Table 3. For the present study, the last information retrieval took place on April 25, 2024.

From the initial search in the first three scientific databases, we retrieved 756 records: 462 from Scopus, 163 from ProQuest, and 131 from Web of Science. We considered this amount to be a small number of records for an SLR. Consequently, we expanded the search to Google Scholar, which indexes both peer-reviewed and non-peer-reviewed sources. To guarantee scientific coherence, we compared non-peer-reviewed records to peer-reviewed ones. On that basis, we included only those that proved coherent with peer-reviewed ones.

Initially, Google Scholar returned over 26,000 records. However, we observed a sharp decline in relevance after the first 200 results, which we sorted by relevance. Thus, to maintain feasibility and ensure quality, we limited our review to the first 250 records from Google Scholar.

As a result, we included 1006 records for the present study, representing a robust and diverse dataset for addressing the research objectives.

TABLE 1 | Keywords for Boolean search string generation.

Large language models	Cognitive assistants	Manufacturing
LLM	Conversational AI	Quality management systems
Conversational AI Models	AI assistant	Industry 4.0
Generative Pre-trained Transformer models	Machine learning agent	Fabrication
GPT	Assistant	Processing
Generative AI	Chatbot	Production
GenAI		
Generative Artificial Intelligence Models		

3.1.3 | Selection Process

We followed a multi-stage selection process:

- We removed 161 duplicates, leaving 845 records;
- We screened titles and abstracts using inclusion criteria three (3) and four (4), which led us to select 83 reports for full-text retrieval;
- We excluded 8 reports because of access restrictions, leaving 75 articles for full-text analysis;
- We excluded 22 studies after full-text evaluation, resulting in 53 studies included in the final review.

Figure 3 illustrates the selection process and outlines the reasons for exclusion.

The first author of this study performed the screening and selection independently. In cases of uncertainty, the second author provided input. Together, we reached consensus on all inclusion and exclusion criteria, ensuring a rigorous and transparent review process.

3.1.4 | Data Collection Process

Before initiating the review process, we identified key terms related to challenges associated with LLM-CAs. To capture a broader range of relevant language, we included terms such as

TABLE 2 | Inclusion and exclusion criteria with justification.

Inclusion criteria	Justification	Exclusion criteria
1. Be published in English or German	Comprises the languages spoken by the research team	Published in languages other than English or German
2. Be published between 2014 and 2024	In 2014, Goodfellow [62] coined the term Generative Adversarial Network, basic for the further development of Generative AI (GenAI)	Outside the mentioned time span
2. Clearly states interactions between GenAI-cognitive assistants and humans at work	Comprises the main scope of investigation	No GenAI-cognitive assists and human interactions
3. Be related or applicable to manufacturing	Manufacturing is the environment of investigation	Not relatable to manufacturing

TABLE 3 | Selection of information sources.

Information source	Justification	Filters
Scopus	Indexes a vast range of engineering sources, for example, journals, conferences, often overlapping with other disciplines, for example, computer science.	*Source type*: conference, article, review, conference review, book chapter, book.
ProQuest	Provides access to a variety of sources in the engineering field, for example, journals. Moreover, it includes its own ProQuest engineering collection, which focuses on applied and theoretical engineering aspects.	*Source type*: books, dissertations and theses, magazines, professional magazines, scientific journals. *Document type*: article, report, book, book chapter, business case, dissertation and thesis, case study.
Web of Science	Indexes rigorous, high-impact journals in engineering, offering a broad interdisciplinarity.	*Source type*: article, review article, proceeding paper, early access, editorial material.
Google Scholar	In the scope of this study, it offers an extension to the previous three databases, once the topic investigated is new. Therefore, literature such as work papers is also included with the goal of obtaining a holistic view of current research in the field.	No filter. Sorted by relevance. Limited to the first 250 records.

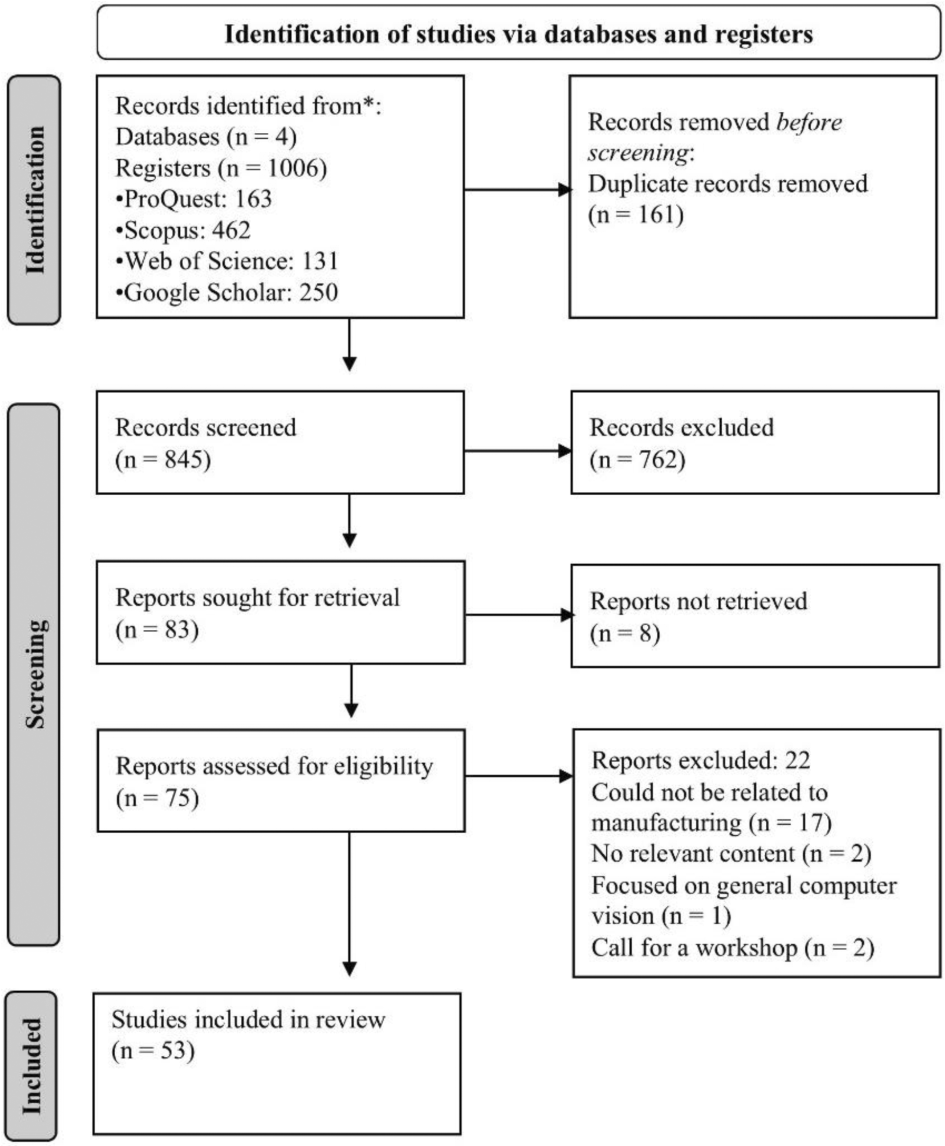

FIGURE 3 | Selection process of studies—PRISMA flow diagram.

TABLE 4 | Example of a database entry.

Title	Authors	Year	Source	Challenges
Review of Generative Artificial Intelligence Use Cases Applicable to Manufacturing Industry	Kulkarni, N. D. Saurav, B.	2024	Journal	Biased content, …

TABLE 5 | Author-challenges matrix.

Title	Authors	Year	Challenge 1	Challenge 2	Challenge 3
Title 1 …	Author 1	Year	x		x
Title 2 …	Author 2	Year		x	

needs, *issues*, *problems*, *obstacles*, *difficulties*, *hindrance*, *complications*, *limitations*, *and requirements*. Throughout this paper, we consistently refer to these terms as "challenges" for clarity and simplicity.

We conducted full-text analysis of all included studies using these challenge-related terms. We managed references and organized knowledge from the selected studies using the Citavi software. Moreover, we saved the identified challenges in a database and correlated each entry to its corresponding study, as exemplified by Table 4.

Most of the challenges appeared in specific sections of the selected studies, that is, introduction, results, discussions, conclusions, and outlook.

The first author reviewed all 53 full-text studies and extracted relevant data. The second author subsequently reviewed this data for accuracy. Both authors resolved discrepancies through consensus, ensuring rigor and consistency in the collection process.

We subdivided the studies into two categories:

- 1st category—studies directly related to manufacturing, for example, [63].
- 2nd category—studies from which results could be transferred to manufacturing, for example, [64].

This categorization allowed us to include a broader set of studies that were not exclusively industry-related to broaden our understanding related to LLM-CAs for QMS. During the review process, we noticed that different authors used varied terminology to describe similar challenges. The challenges "lack of data privacy" and "insufficient data privacy" illustrate this similarity. In such cases, we labeled both terms as "lack of data privacy." We used the same approach for all other identified challenges for standardization.

For each study, we documented authorship, publication year, and source, along with the study's objectives, methods, findings, discussion, conclusions, and outlook. We compiled this data into Table 4, extending it to a larger dataset from the 53 studies.

In sequence, we linked authors to challenges in the form of a matrix, as per Table 5, in which "x" means that an author mentioned a specific challenge. Section 4 presents the full results.

We used the author-challenges matrix as input for the thematic analysis, described in Section 3.2. Two researchers independently grouped challenges based on similarity. After this individual evaluation, both researchers grouped the identified challenges by consensus. After comparing and discussing their individual groupings, the researchers reached consensus on a final structure. We labeled the resulting groups with unique identifiers ("ID" in Section 4), which formed the basis for deriving LLM-CA requirements.

3.2 | Thematic Analysis (TA)

Thematic analysis is a research method for identifying, analyzing, and reporting patterns of meaning within qualitative data [58]. On that basis, a qualitative dataset serves as a source of insights, which is then translated into patterns of meaning, that is, themes [65]. Braun and Clarke [58] state that TA offers both theoretical and design flexibility, allowing researchers to adapt it to research scenarios. Historically, TA evolved to become a widely recognized research method, providing an accessible and flexible method for qualitative data analysis [65].

Braun et al. [66] outline two primary approaches to TA, that is, inductive and deductive. The inductive approach derives themes directly from data without the use of preexisting theories or frameworks. Therefore, it is a bottom-up approach that allows themes to emerge organically, ensuring that the TA remains closely tied to the collected qualitative data. Contrarily, a top-down approach, based on predetermined theories or frameworks, underlies the deductive approach. In this case, existing theories or frameworks guide the TA from a higher level, for example, from concepts to an elementary level, for example, qualitative statements. A combination of both inductive and deductive approaches is possible. Thus, TA can follow some data familiarization (inductive) and derivation from prior theories (deductive) [58].

According to Braun et al. [66], TA follows basically six phases, as per Table 6. Naeem et al. [67] add explicitly an additional step,

TABLE 6 | The six phases of thematic analysis (adapted from [66]).

Phase	Process
1. Familiarization	Getting to know the data through reading and re-reading, and, additionally, identifying keywords.
2. Coding	Tagging data that capture relevant features to the research question.
3. Initial theme generation	Clustering and re-clustering of similar codes aiming at unifying patterns of meaning around a core concept.
4. Reviewing and developing themes	Reviewing and revisioning phase 3 to ensure the aggregation of meaningful patterns, considering both the codes and the entire dataset. It focusses on the story each theme tells.
5. Refining, defining and naming themes	Refining themes through short theme definitions. Researchers name each theme.
6. Producing the report	Writing phase putting together data extracted and the story told by each theme in context of existing knowledge, reflecting on the methodological process.

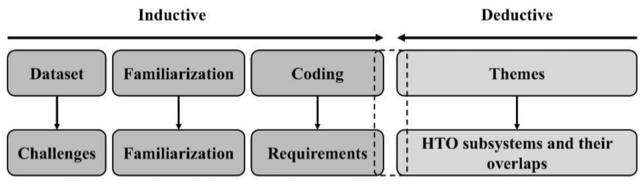

FIGURE 4 | Combination of deductive and inductive approaches.

that is, selection of keywords. While identifying recurring patterns in the qualitative dataset, authors can designate them as keywords, which capture perceptions from the dataset. We incorporated this step into Phase 1 of Braun et al. [66] to support pattern recognition and coding.

We followed a combination of both deductive and inductive approaches. In the deductive component, we predefined the themes based on the HTO concept, including both independent and overlapping subsystems. Consequently, the predefined themes are:

1. Human
2. Technology
3. Organization
4. Human and technology
5. Human and organization
6. Organization and technology
7. Human, technology and organization

In the inductive component, we used the identified challenges, as per Table 5, as our qualitative dataset. We extracted keywords and coded them based on recurrence and similarity. The interface between these codes and the predefined themes served as the foundation for the thematic analysis. Figure 4 depicts the integration of deductive and inductive approaches. First, we derived a dataset from the literature. Second, we translated these challenges into LLM-CAs' requirements for QMS, defining the coding.

To finalize the requirements, we conducted three collaborative research group meetings, aiming at reaching consensus. In the first meeting, both authors independently categorized the challenges without assigning them to themes. In the second, we compared and refined these preliminary categorizations. Thereafter, we drafted a shared categorization scheme and converted each group into a requirement, generating the first consensual draft. In the final group meeting, we revised the structure and wording of the requirements until we reached consensus. Finally, we aligned requirements (codes) with the matching themes derived from the HTO concept through consensus.

This combined approach enabled us to develop a comprehensive requirement analysis for deploying LLM-CAs in manufacturing under QMS constraints. We present the complete results in Section 4.

4 | Results

This section presents the findings of the study. First, we include a descriptive analysis of the SLR. Second, we identify challenges related to LLM-CAs for QMS, answering the first sub-research question. Finally, we answer the second sub-research question through the derived requirements and their classification in accordance with the HTO concept.

4.1 | Descriptive Analysis of the SLR

The SLR comprised a full-text analysis of 53 studies (Figure 3). These studies, published between 2022 and 2024, came from various sources, that is, scientific journals, conferences, and scientific repositories. As a result, our dataset comprises journal papers, conference papers, working papers, and preprints. Tables 7–9 exhibit the descriptive breakdown of the included studies.

We observed a sharp increase in publication volume between 2022 (1 study) and 2023 (36 studies). This growth reflects the academic response to OpenAI's release of ChatGPT 3.5 in November 2022 [24], which triggered academic interest in LLMs. Publications from 2023 represent 68% of our dataset, highlighting the centrality of that year to current research. By April 2024, studies

TABLE 7 | Number and percentage of included studies per year.

Year of publication	Number of articles ($N = 53$)	Percentage (%)	Cumulative percentage (%)
2022	1	2.0	2.0
2023	36	68.0	70.0
2024 (January–April)	16	30.0	100.0

TABLE 8 | Number and percentage of included studies per source of publication.

Source of publication	Number of articles ($N = 53$)	Percentage (%)	Cumulative percentage (%)
Journal	31	59.0	59.0
Conference	16	30.0	89.0
Scientific repository	6	11.0	100.0

TABLE 9 | Number and percentage of included studies per type of publication.

Type of publication	Number of articles ($N = 53$)	Percentage (%)	Cumulative percentage (%)
Journal article	31	59.0	59.0
Conference article	16	30.0	89.0
Working paper	3	5.5	94.5
Preprint	3	5.5	100.0

from that year accounted for 30% of the total number of articles included, showing continued and accelerating relevance.

Although we included Google Scholar as an information source, most of its studies originated from scientific journals and conferences. These accounted for 59.0% and 30.0%, respectively. The remaining 11.0% consisted of working papers and preprints from scientific repositories.

Table 10 presents the 84 unique challenges we identified across the 53 studies. We grouped similar challenges under shared IDs, which link directly to the requirements presented in Section 4.2 "Findings from the TA." Accordingly, we answered the first sub-research question, which had the goal of identifying literature-based challenges related to LLM-CAs for QMS.

Several challenges emerged as particularly prominent because of their frequent citation by multiple authors. The following challenges appeared in at least 10 studies:

- Lack of data privacy/security (15 authors).
- Hallucination (14 authors).
- Resource-intensive integration and implementation of AI systems (14 authors).
- Lack of interpretability and transparency in the decision-making process—black box (11 authors).
- Biased content (10 authors).
- Lack of contextual awareness (10 authors).
- Lack of skilled workforce who can use, develop, manage, and maintain AI systems (10 authors).

These frequent citations emphasize their relevance regarding LLM-CAs for QMS. Notably, the highly cited challenges are technological ones. While advancing technology can potentially address many of these challenges, it alone is not sufficient. Other significant challenges relate to human and organizational dimensions such as variations in accents and dialects, and change management necessity, respectively. Addressing human and organizational challenges is time-consuming and demands strategic engagement over the short, medium, and long-term.

Finally, Table 11 summarizes the research methods used to identify the challenges. Most of the authors relied on case studies and non-systematic literature reviews, while only a few employed mixed-method approaches.

4.2 | Findings From the TA

After grouping similar challenges, we derived a set of requirements for LLM-CAs for QMS. Table 12 presents the complete list of requirements, which we classified according to the HTO subsystems and their overlaps.

The TA produced 24 distinct requirements (codes) distributed across seven (7) themes, each aligned with the HTO concept. The TA reveals that most requirements fall within the "technology" subsystem (9 codes) and the "human-technology" overlap (6 codes). In contrast, we identified only one requirement each under the "human" subsystem and the "human-organization" overlap.

In summary, these findings suggest that the successful deployment of LLM-CAs for QMS relies heavily on technological development and on human–technology interaction. At the same time, the organizational subsystem plays a pivotal role.

Two requirements fall exclusively under the organization subsystem, that is, "implement comprehensive organizational governance" and "conduct empirical studies in Operations & Supply Chain Management (O&SCM)". Moreover, we identified six further requirements that span across subsystems, that is, "human-organization," "technology-organization," and "human-technology-organization."

Together, these findings underscore the need for an integrated, system-level approach to LLM-CAs for QMS. This approach must, simultaneously, address human, technological, and organizational subsystems in line with the HTO concept.

5 | Discussion

The integration of LLM-CAs for QMS demands close alignment with both ISO 9001:2015 clauses and the HTO subsystems. This section discusses how the 24 derived requirements map onto the seven ISO 9001:2015 clauses. On that basis, we answer the third sub-research question, demonstrating how LLM-CAs can support QMS, especially in the context of digital transformation in manufacturing.

TABLE 10 | Identified challenges ($N = 84$, ID = 24).

ID	Challenge
1	Adaptation to noisy environments and adverse conditions in factories [7]
2	Adversarial prompts [69]
	Generation of incorrect answers in face of misleading or incorrect inputs [10, 11, 15, 70, 71]
3	Complex prompting [14, 73–75]
	Limited context capabilities through prompting [78]
4	Unethical usage [19, 20, 32, 64, 79]
	Potential plagiarism [80]
	Copyright-infringed content [64, 82]
	Lack of data privacy/security [7]
	Non-compliance with industry regulations or legal standards [12, 32, 82]
	Privacy invasion [4]
5	Dependency on high quality data [20, 32, 80, 82, 84, 85]
	Difficulty obtaining data to train the models [83]
	No optimized data usage [10]
	Need of manual/human data curation [85, 87]
6	Biased content [13, 15, 19, 32, 64, 68, 69, 79, 81, 82]
	Biases in training data leads to reinforcement of harmful social norms [12, 70, 81, 82]
7	Complexity of domain-specific vocabulary [63]
	Lack of business context understanding [11]
	Lack of NLP algorithms customized for the manufacturing domain [91]
	Lack of contextual awareness [11, 15, 20, 32, 70, 76, 78, 86, 90, 93]
	Lack of domain-specific knowledge [7, 14, 72, 77, 91, 94, 95]
8	Challenging generalization and adaptability [7, 14, 32, 79, 81, 82, 84]
	Limited flexibility [91]
	Scaling challenges for AI Systems in worldwide applications [79]
9	Hallucination [4, 10, 11, 14, 15, 18–20, 64, 69, 76, 77, 96, 97]
	Content disparity compared to a control dataset [98]
	Reasoning errors [29]
	Unreliable and inaccurate output [12, 14, 19, 23, 29, 32, 75, 80, 95]
	Misinformation [64, 69]
	No guarantees on recommendations or correctness of explanations [4]
	Undue confidence in responses without indicating uncertainty levels [32]
	Limited output dialog size [94]
	Incomplete answers triggered by long prompts [74]
	Unsuitable for precise or quantified decision-making problems involving extensive, high-dimensional information [93]
	Limited precision and recall in data transformation tasks [18]
	Performance Variation with Document Size [96]
	Causation aversion [31]
	Handling ambiguity and uncertainty [72]
	Output variability for the same query [73, 74, 96, 102]
	Limited model's parameter capacity [100]
	Lack of interpretability and transparency in the decision-making process—black box [10, 12–14, 19, 32, 64, 80, 83, 102]
10	Model forgetting because of prolonged training on generated data over long run [68]
	Lack of memory [20]
11	Limited real-time responsiveness [7, 72]
	System instability leading to crashes [76, 77]
12	Insufficient evaluation metrics for GenAI Chatbots [32, 68]
	Necessity of verification of AI-generated reports/information [12, 14, 68, 74]
13	Frequent necessity of model update [14, 32, 80, 81]
	Restricted ability to utilize new data (after last training) [10]
	Model's training dependability [80, 83]
14	Ambiguous and context-dependent language used by operators [63]
	Limited subtlety and nuance comprehension in human language [32]
	Variations in accents and dialects [63]
	Variations in accents, intonations, and speech patterns [86]
15	Semantic mistakes [32, 74]
	Grammar mistakes [32]
16	Lack of a skilled workforce who can use, develop, manage, and maintain AI systems [15–17, 72, 79, 82–85, 88]
	Insufficient knowledge on human–AI interaction [89]
	Need of interdisciplinary teams of researchers and practitioners from technical, social, economic, and ethical disciplines [23]
17	Lack of user feedback [15, 64, 73, 90]
	Lack of tacit knowledge [92]
	Lack of incorporated human intuition [31]
	Lack of feedback loops [75]
18	Inability to experience emotions or empathy [32, 63, 70]
19	Low user friendliness [91]
	Need for User Interface Enhancement [63, 96]
	User safety and well-being [20]
	Limitation in Interpreting Visual Elements in PDF files [96]
	Multimodal human–ChatBot interaction [64, 95, 99]
20	Risk of Monopolistic Power and Inefficiency in Decision Making [100]
	Change management necessity [79]
	Clear accountability for errors or accidents caused by AI-generated outputs [14]
	Lack of compensation policy for workers who provide data to AI [92]
	No cost–benefit analysis [11, 12]
21	Mistakes in decision-making in face of miscommunication or inaccurate communication [99]
22	Lack of empirical studies in Operations and Supply Chain Management (O&SCM) [101]
23	AI avoidance [16, 32, 64]
	Low client confidence/trust [11, 16, 69]
	AI overreliance and scapegoating [16, 20, 32, 64]
24	Consolidation and validation difficulties of AI applications because of data volume created across supply chain nodes [79]
	Multiple dataset integration [100]
	Difficult integration with legacy and heterogeneous systems in factories [7]
	Resource-intensive integration and implementation of AI systems [11, 12, 14, 15, 17, 72, 79, 82–86, 88, 91]

TABLE 11 | Applied research methods and number of authors.

Methods	Number of authors
Case study	25
Literature review (non-systematic)	14
Systematic literature review	5
Experiment and survey	3
Survey	3
Interviews and qualitative analysis	1
Literature review, proof of concept, and focus groups	1
Literature review and taxonomy development model	1

TABLE 12 | Classification of requirements in line with the HTO subsystems and their overlaps.

ID	Requirements (coding)	HTO subsystems and their overlaps (themes)
23	Enable AI trustworthiness	Human
12	Evaluate LLM-CAs performance	Technology
2	Improve resistance to adversarial input	
8	Enhance adaptability and scalability	
24	Enable seamless data and systems harmonization	
9	Ensure reliable model performance	
10	Guarantee effective memory usage	
11	Guarantee stable system performance	
13	Enable reliable training and update of LLMs	
15	Ensure accurate language output	
20	Implement comprehensive organizational governance	Organization
22	Conduct empirical studies in O&SCM	
3	Comprise flexible prompt capabilities	Human-technology
14	Ensure robust language understanding	
17	Include human-in-the-loop	
18	Integrate emotional intelligence	
19	Design user-centric cognitive assistants	
21	Enable effective communication-driven decision-making	
16	Develop a skilled workforce	Human-organization
1	Adapt to factory environments	Technology-organization
7	Integrate industry-specific knowledge	
6	Mitigate biased content	Human-technology-organization
4	Ensure compliance integrity	
5	Guarantee robust data quality management	

5.1 | Clause "Context of the Organization"

This clause requires organizations to identify the needs and expectations of stakeholders, for example, customers, suppliers, and regulators, and define the scope of their QMS. Additionally, organizations must establish and maintain processes that align with their strategic direction and ensure continuous improvement [8].

We mapped the following requirements to this clause:

- ID 7: Integrate industry-specific knowledge.
- ID 20: Implement comprehensive organizational governance.
- ID 22: Conduct empirical studies in operations and supply chain management (O&SCM).

These requirements ensure that LLM-CAs are developed with a deep understanding of industry-specific needs and are governed appropriately within strategic and contextual boundaries.

5.2 | Clause “Leadership”

ISO 9001:2015 emphasizes leadership accountability in establishing clear quality policies, in alignment with strategic organizational goals. Moreover, leaders must ensure that roles, responsibilities, and authorities are well defined. They must also promote a culture of customer focus, continuous improvement, and effective communication throughout the organization [8].

LLM-CA requirements aligned with this clause include:

- ID 19: Design user-centric cognitive assistants.
- ID 21: Enable communication-driven decision-making.
- ID 4: Ensure compliance integrity.

These requirements underscore the role of leadership in ensuring that LLM-CAs align with organizational ethics, decision-making values, and communication practices.

5.3 | Clause “Planning”

Organizations shall identify and address risks and opportunities that could impact their QMS. It also mandates setting measurable quality objectives and planning actions to achieve them. Moreover, it requires managing changes systematically to ensure continuous improvement [8].

Mapped LLM-CA requirements are:

- ID 23: Enable AI trustworthiness.
- ID 2: Improve resistance to adversarial input.
- ID 6: Mitigate biased content.
- ID 13: Enable reliable training and update of LLMs.
- ID 8: Enhance adaptability and scalability.

These address risk management, system resilience, and the ability to evolve LLM-CAs in line with changing operational needs.

5.4 | Clause “Support”

This clause focuses on providing resources, competence, and infrastructure to maintain an effective QMS. It also emphasizes the importance of awareness, communication, and proper documentation to ensure consistent implementation and control of quality processes [8].

Aligned requirements include:

- ID 1: Adapt to factory environments.
- ID 16: Develop a skilled workforce.
- ID 14: Ensure robust language understanding.
- ID 10: Guarantee effective memory usage.
- ID 18: Integrate emotional intelligence.
- ID 17: Include human-in-the-loop.

Together, these requirements emphasize the human–technology interface and the importance of equipping workers and systems to interact meaningfully.

5.5 | Clause “Operation”

This clause covers the execution of processes needed to meet customer requirements, including planning, production, service provision, and supplier control. It also addresses risk management, change control, and ensuring that products and services conform to specified quality standards throughout their lifecycle [8].

Relevant LLM-CA requirements:

- ID 15: Ensure accurate language output.
- ID 3: Comprise flexible prompt capabilities.
- ID 24: Enable seamless data and systems harmonization.

These capabilities allow LLM-CAs to support operational decision-making and real-time shop floor interactions efficiently and flexibly.

5.6 | Clause “Performance Evaluation”

ISO 9001:2015 requires organizations to monitor, measure, analyze, and evaluate the performance of their QMS. This includes conducting internal audits, reviewing customer satisfaction, and assessing the effectiveness of the QMS to ensure continuous improvement and compliance with objectives [8].

Mapped LLM-CA requirements:

- ID 12: Evaluate LLM-CAs performance.
- ID 9: Ensure reliable model performance.
- ID 11: Guarantee stable system performance.

These requirements support continuous monitoring, auditing, data-driven refinement of LLM-CAs, and system stability in real-world manufacturing contexts.

5.7 | Clause “Improvement”

This clause focuses on organizations’ need to correct identified nonconformities and to continuously improve the effectiveness of their QMS. It encourages the use of data and feedback to drive improvements and ensure ongoing compliance with quality objectives [8].

Mapped requirements:

- ID 5: Guarantee robust data quality management.

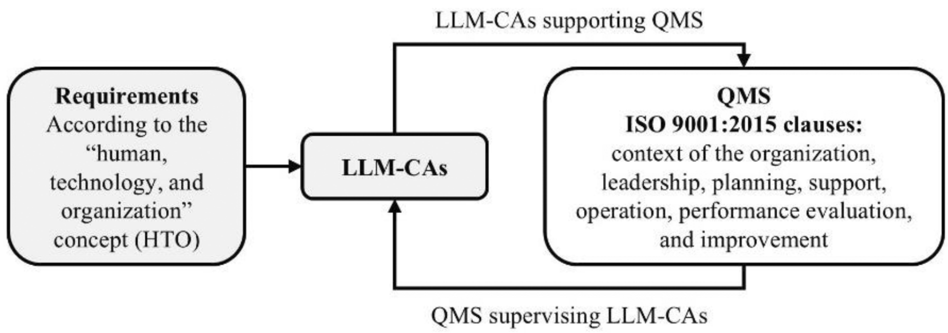

FIGURE 5 | Interplay between LLM-CAs and QMS.

- ID 15: Ensure accurate language output.
- ID 13: Enable reliable training and update of LLMs.
- ID 17: Include human-in-the-loop.

This highlights LLM-CAs' potential to support employee-driven, iterative process improvement, particularly when humans remain actively involved in system learning loops.

5.8 | Summary

Figure 5 illustrates a strong interplay between LLM-CAs and ISO 9001:2015 clauses. This interplay emphasizes that the successful integration of LLM-CAs for QMS is not purely a technical endeavor. Integration requires a systemic transformation requiring alignment across human, technological, and organizational dimensions. This mapping reveals that LLM-CAs also serve as enablers of QMS evolution, particularly in Industry 4.0 and are at the same time subjected to QMS.

5.9 | Empirical Illustration of Requirements

The following hypothetical LLM-CA for QMS case study serves as a basis for making the study more accessible to practitioners and strengthening empirical grounding.

5.9.1 | Optimizing Product Setup Times in a Manufacturing Line

Setup time optimization is a persistent challenge in manufacturing, particularly in multi-product environments. Each time new products are added to a production line, planners must minimize changeover times while maintaining product quality and workflow continuity. LLM-CAs for QMS can support this process.

Consider a mid-sized manufacturer in the automotive sector that routinely introduces new products into its production schedule. Integrating new products requires re-sequencing tasks, considering factors such as tooling requirements, inspection needs, shift schedules, and defect rates. This process relies, traditionally, on expert intuition and expertise, and siloed data from manufacturing systems and QMS, resulting in inefficiencies.

The company can implement an LLM-CA to overcome this issue. The assistant simulates and recommends optimized production sequences by analyzing structured and unstructured data across systems. On that basis, it generates human-readable rationales for each recommendation, supporting real-time human input to adjust sequencing decisions. Finally, it documents the decision-making process for auditability and process learning.

The assistant serves, therefore, as an interface between planning, operations, and quality management. Functioning this way, it ensures that any sequencing plan aligns with strategic objectives and complies with process constraints. Therefore, integrating this LLM-CA directly supports several ISO 9001:2015 clauses and reflects specific HTO-aligned requirements derived from this study. Exemplarily, the assistant supports the clause "operation," in which it assesses data from diverse sources, enabling seamless data and systems harmonization. Moreover, it integrates industry-specific knowledge, supporting the "clause context of the organization," leading consequently to continuous improvement.

5.10 | Study Limitations

Although providing a structured investigation of challenges related to the integration of LLM-CAs for QMS, this study faces several limitations. First, we identified challenges using only one scientific method, that is, SLR. Therefore, we excluded other perspectives, for example, opinions of experts or shop floor workers. On that basis, practical challenges are not exhaustively covered. Second, although the search period spanned from 2014 to 2024, we considered only articles published between 2022 and 2024. This is primarily because of the release of ChatGPT-3.5 by OpenAI in November 2022 [24], which significantly intensified academic interest in LLMs. As a result, 68% of the studies are from 2023, with another 30% from early 2024, that is, January to April. This distribution indicates a continued and growing research interest in the topic. However, it also introduces a publication cycle bias, as many 2024 studies may not yet be published or indexed.

6 | Conclusion and Outlook

This study proposed a holistic and organized classification of literature-based requirements for LLM-CAs for QMS, as per ISO

9001:2015, according to the HTO concept. Through an SLR of 53 studies published between 2022 and 2024, we identified 84 unique challenges, which we then translated into 24 actionable requirements using TA.

Grounded in the HTO concept, we categorized these requirements according to each HTO subsystem and their overlaps. Our analysis reveals that most of the requirements lie within the "technology" and "human-technology" domains. Accordingly, there is a need for advanced LLM-CA capabilities and effective human-LLM-CA interaction. However, the "organization" subsystems also emerged as essential for enabling structured governance and long-term adoption.

By mapping the derived requirements to the seven ISO 9001:2015 clauses—context of the organization, leadership, planning, support, operation, performance evaluation, and improvement—we showed a strong alignment between LLM-CA integration and established QMS principles. Requirements such as "enable AI trustworthiness," "mitigate biased content," "include human-in-the-loop," and "evaluate LLM-CAs performance" are not only technically relevant but also contribute to compliance, risk reduction, and operational excellence within QMS.

To enhance practical relevance, we illustrated the integration of LLM-CAs through a hypothetical case study focused on setup time optimization in a multi-product manufacturing line. This real-world application applied key requirements and showed how LLM-CAs can support decision-making, foster continuous improvement, and align with QMS objectives.

The technical realization of the derived requirements is a challenging task. Future research should focus, at least, on their partial implementation. Based on that, integrative solutions could be enabled once LLM technologies evolve. For instance, a fully integrated solution could enable shop floor workers to improve their own knowledge base through digital feedback loops. Additionally, data analysis components could be combined with LLM-CAs, which could assist workers in performing continuous process improvement.

Future research should incorporate other scientific methods, for example, focus groups, interviews, surveys, and empirical testing to refine and validate derived requirements. Moreover, a thorough integration of stakeholders, for example, managers, shop floor workers, and regulators, can contribute to a broader scientific approach. Furthermore, designing a modular software architecture based on these requirements will help operationalize LLM-CAs for QMS in real settings. Once proven deployable, such a solution must also account for critical human and organizational dimensions. These future developments are part of our ongoing work and will be presented to the scientific community as they progress.

In sum, this study offers a foundation for integrating LLM-CAs for QMS by aligning AI capabilities with ISO standards and human-centered design. This alignment is critical for ensuring trust, transparency, and sustainable AI adoption in Industry 4.0 environments.

Author Contributions

Marcos Galdino: conceptualization, data curation, formal analysis, investigation, methodology, writing – original draft. **Tobias Hamann:** methodology, validation, writing – review and editing. **Anas Abdelrazeq:** supervision, validation, writing – review and editing. **Ingrid Isenhardt:** supervision, writing – review and editing.

Acknowledgments

This research was developed in the context of the project ZUKIPRO [ZUK-1-0005], which is funded by the German Federal Ministry of Labour and Social Affairs (BMAS) and the European Union via the European Social Fund Plus (ESF Plus) as part of the "Future Centres" program. Open Access funding enabled and organized by Projekt DEAL.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

1. D. Chen, Y. Huang, Z. Ma, et al., "Data-Juicer: A One-Stop Data Processing System for Large Language Models," in *Companion of the 2024 International Conference on Management of Data* (The Association for Computing Machinery, 2024), 120–134.

2. M. Jin, S. Wang, L. Ma, et al., "Time-LLM: Time Series Forecasting by Reprogramming Large Language Models," in *International Conference on Representation Learning* (2024), 23857–23880, https://proceedings.iclr.cc/paper_files/paper/2024/hash/680b2a8135b9c71278a09cafb605869e-Abstract-Conference.html.

3. A. D. Paroha and A. Chotrani, "A Comparative Analysis of TimeGPT and Time-LLM in Predicting ESP Maintenance Needs in the Oil and Gas Sector," *International Journal of Computer Applications* 168, no. 8 (February 2024): 975–8887.

4. S. Kernan Freire, M. Foosherian, C. Wang, and E. Niforatos, *Harnessing Large Language Models for Cognitive Assistants in Factories* (ACM, 2023), 1–6.

5. M. Elg, A. Birch-Jensen, I. Gremyr, J. Martin, and U. Melin, "Digitalisation and Quality Management: Problems and Prospects," *Production Planning & Control* 32, no. 12 (2020): 990–1003.

6. R. H. Todd, W. E. Red, S. P. Magleby, and S. Coe, "Manufacturing: A Strategic Opportunity for Engineering Education," *Journal of Engineering Education* 90 (2001): 397–405.

7. R. Pereira, C. Lima, T. Pinto, and A. Reis, "Virtual Assistants in Industry 4.0: A Systematic Literature Review," *Electronics* 12, no. 19 (2023): 4096.

8. International Organization for Standardization, *ISO 9000:2015 (E): Quality Management Systems - Requirements* (Switzerland, 2015).

9. S. M. Lee, D. Lee, and Y. S. Kim, "The Quality Management Ecosystem for Predictive Maintenance in the Industry 4.0 Era," *International Journal of Quality Innovation* 5, no. 1 (2019): 4.

10. S. Ghodratnama and M. Zakershahrak, *Adapting LLMs for Efficient, Personalized Information Retrieval: Methods and Implications*, vol. 14518 (Springer Nature Singapore, 2024), 364.

11. R. Raj, A. Singh, V. Kumar, and P. Verma, "Analyzing the Potential Benefits and Use Cases of ChatGPT as a Tool for Improving the Efficiency and Effectiveness of Business Operations," *BenchCouncil Transactions on Benchmarks, Standards and Evaluations* 3, no. 3 (2023): 100140.

12. N. Rane, "ChatGPT and Similar Generative Artificial Intelligence (AI) for Smart Industry: Role, Challenges and Opportunities for Industry 4.0, Industry 5.0 and Society 5.0," *Innovations in Business and Strategic Management* 2, no. 1 (2024): 10–17.

13. A. Ainapure, S. Dhamane, and S. Dhage, "Embodied Epistemology: A Meta-Cognitive Exploration of Chatbot-Enabled Document Analysis," in *2023 International Conference on Evolutionary Algorithms and Soft Computing Techniques (EASCT)* (2023), 1–6, https://doi.org/10.1109/EASCT59475.2023.10392618.

14. P. Ghimire, K. Kim, and M. Acharya, "Opportunities and Challenges of Generative AI in Construction Industry: Focusing on Adoption of Text-Based Models," *Buildings* 14, no. 1 (2024): 220.

15. J. Bandlamudi, K. Mukherjee, P. Agarwal, et al., "Building Conversational Artifacts to Enable Digital Assistant for APIs and RPAs," in *Proceedings of the AAAI Conference on Artificial Intelligence* 38, no. 21 (2024): 22725–22733, https://doi.org/10.1609/aaai.v38i21.30306.

16. E. Bendoly, A. Chandrasekaran, M. D. Lima, R. Handfield, S. H. Khajavi, and S. Roscoe, "The Role of Generative Design and Additive Manufacturing Capabilities in Developing Human–AI Symbiosis: Evidence From Multiple Case Studies," *Decision Sciences* 55, no. 4 (2023): 325–345, https://doi.org/10.1111/deci.12619.

17. N. D. Kulkarni Saurav, "Review of Generative Artificial Intelligence Use Cases Applicable to Manufacturing Industry," *International Journal of Science and Research (IJSR)* 13, no. 1 (2024): 1189–1192.

18. A. Berenguer, A. Morejón, D. Tomás, and J. N. Mazón, "Using Large Language Models to Enhance the Reusability of Sensor Data," *Sensors* 24, no. 2 (2024): 347.

19. Q. Ding, D. Ding, Y. Wang, C. Guan, and B. Ding, "Unraveling the Landscape of Large Language Models: a Systematic Review and Future Perspectives," *Journal of Electronic Business & Digital Economics* 3, no. 1 (2024): 3–19.

20. V. Hassija, A. Chakrabarti, A. Singh, V. Chamola, and B. Sikdar, "Unleashing the Potential of Conversational AI: Amplifying Chat-GPT'S Capabilities and Tackling Technical Hurdles," *IEEE Access* 11 (2023): 143657–143682.

21. M. Berglund, A. Karltun, J. Eklund, and J. Karltun, "HTO—A Concept of Humans, Technology and Organization in Interaction," *HELIX Working Papers* 20 (2020): 2.

22. A. Karltun, J. Karltun, M. Berglund, and J. Eklund, "HTO - A Complementary Ergonomics Approach," *Applied Ergonomics* 59, no. Pt A (2017): 182–190.

23. D. C. Doanh, Z. Dufek, J. Ejdys, et al., "Generative AI in the Manufacturing Process: Theoretical Considerations," *Engineering Management in Production and Services* 15, no. 4 (2023): 76–89.

24. W. X. Zhao, K. Zhou, J. Li, et al., "A Survey of Large Language Models," *arXiv* 1 (2024): 140.

25. A. Vaswani, N. Shazeer, N. Parmar, et al., "Attention Is All You Need," in *31st Conference on Neural Information Processing Systems (NIPS)* (USA, 2017).

26. Open AI, "GPT-4 Technical Report," 2023.

27. T. B. Brown, B. Mann, N. Ryder, et al., "Language Models Are Few-Shot Learners," *Advances in Neural Information Processing Systems* 33 (July 2020): 1877–1901.

28. X. Li, S. Yuan, X. Gu, Y. Chen, and B. Shen, "Few-Shot Code Translation via Task-Adapted Prompt Learning," *Journal of Systems and Software* 212 (2024): 112002.

29. S. Kernan Freire, *LLM-Powered Cognitive Assistants for Knowledge Sharing Among Factory Operators* (Delft University of Technology, 2025).

30. R. Figliè, T. Turchi, G. Baldi, and D. Mazzei, "Towards an LLM-Based Intelligent Assistant for Industry 5.0," in *Proceedings of the 1st International Workshop on Designing and Building Hybrid Human–AISystems (SYNERGY2024)* (CEUR Workshop Proceedings, 2024).

31. S. Saghafian, "Effective Generative AI: The Human-Algorithm Centaur," 2023.

32. F. Khennouche, Y. Elmir, Y. Himeur, N. Djebari, and A. Amira, "Revolutionizing Generative Pre-Traineds: Insights and Challenges in Deploying ChatGPT and Generative Chatbots for FAQs," *Expert Systems with Applications* 246 (2024): 123224.

33. M. D. Jones, S. Hutcheson, and J. D. Camba, "Past, Present, and Future Barriers to Digital Transformation in Manufacturing: A Review," *Journal of Manufacturing Systems* 60 (2021): 936–948.

34. B. Wang, "The Future of Manufacturing: A New Perspective," *Engineering* 4, no. 5 (2018): 722–728.

35. M. A. Shabur, "A Comprehensive Review on the Impact of Industry 4.0 on the Development of a Sustainable Environment," *Discover Sustainability* 5, no. 1 (2024): 97.

36. B. Dafflon, N. Moalla, and Y. Ouzrout, "The Challenges, Approaches, and Used Techniques of CPS for Manufacturing in Industry 4.0: A Literature Review," *International Journal of Advanced Manufacturing Technology* 113, no. 7–8 (2021): 2395–2412.

37. Z. Suleiman, S. Shaikholla, D. Dikhanbayeva, E. Shehab, and A. Turkyilmaz, "Industry 4.0: Clustering of Concepts and Characteristics," *Cogent Engineering* 9, no. 1 (2022): 2034264.

38. W. E. Forum, *Future of Jobs Report* (World Economic Forum, 2023).

39. S. Banitaan, G. al-refai, S. Almatarneh, and H. Alquran, "A Review on Artificial Intelligence in the Context of Industry 4.0," *International Journal of Advanced Computer Science and Applications* 14, no. 2 (2023): 23–30.

40. L. M. Fonseca, "Industry 4.0 and the Digital Society: Concepts, Dimensions and Envisioned Benefits," *Proceedings of the International Conference on Business Excellence* 12, no. 1 (2018): 386–397.

41. M. Calış Duman and B. Akdemir, "A Study to Determine the Effects of Industry 4.0 Technology Components on Organizational Performance," *Technological Forecasting and Social Change* 167 (2021): 120615.

42. K. Fettig, T. Gačić, A. Köskal, A. Kühn, and F. Stuber, *Impact of Industry 4.0 on Organizational Structures* (IEEE, 2018), 1–8.

43. P. Fantini, M. Pinzone, and M. Taisch, "Placing the Operator at the Centre of Industry 4.0 Design: Modelling and Assessing Human Activities Within Cyber-Physical Systems," *Computers & Industrial Engineering* 139 (2020): 105058.

44. M. Sony, "Pros and Cons of Implementing Industry 4.0 for the Organizations: a Review and Synthesis of Evidence," *Production & Manufacturing Research* 8, no. 1 (2020): 244–272.

45. D. Hoyle, *ISO 9000 Quality Systems Handbook*, 7th ed. (Routledge Taylor & Francis Group, 2018), 1.

46. E. Psomas and A. Pantouvakis, "ISO 9001 Overall Performance Dimensions: an Exploratory Study," *TQM Journal* 27, no. 5 (2015): 519–531.

47. L. Fonseca and J. P. Domingues, "ISO 9001:2015 Edition-Management, Quality and Value," 2017.

48. A. Bakhtiar, A. Nugraha, H. Suliantoro, and D. Pujotomo, "The Effect of Quality Management System (ISO 9001) on Operational Performance of Various Organizations in Indonesia," *Cogent Business & Management* 10, no. 2 (2023), https://doi.org/10.1080/23311975.2023.2203304.

49. N. P. Solomon, A. Bester, and C. M. Moll, "Diffusion of a Quality Management System: a Case Study," *South African Journal of Industrial Engineering* 28, no. 2 (2017): 149–164, https://doi.org/10.7166/28-2-1762.

50. H. Mustroph and S. Rinderle-Ma, "Design of a Quality Management System Based on the EU AI Act," in *Legal Knowledge and Information Systems* (IOS Press, 2024).

51. D. P. Fadilasari, R. Roy Ghatak, J. A. Garza-Reyes, R. Joshi, and J. Kandasamy, "Adopting Quality Management Practices in the Industry 4.0 Era: an Investigation Into the Challenges," *Total Quality Management & Business Excellence* 35, no. 9–10 (2024): 1098–1123.

52. F. Murmura, F. Musso, L. Bravi, and G. Pierli, "The Role of Quality Management Systems in Fostering the International Competitiveness of Companies," *International Journal of Quality & Reliability Management* 41, no. 7 (2024): 1979–1999.

53. A. Karltun, "Developing HTO System's Thinking for Organizational and Technological Change," in *Proceedings of the 10th International Symposium on Human Factors in Organizational Design and Management* (International Ergonomics & Human Factors Association, 2011), 4–6.

54. M. Berglund and J. Karltun, "Human, Technological and Organizational Aspects Influencing the Production Scheduling Process," *International Journal of Production Economics* 110, no. 1–2 (2007): 160–174.

55. A. L. Aas and S. O. Johnsen, *Improvement of Human Factors in Control Centre Design - Experiences Using ISO 11064 in the Norwegian Petroleum Industry and Suggestions for Improvements* (International Petroleum Technology Conference, 2007).

56. K. Olofsson Hallén, M. Forsman, and A. Eriksson, "Interactions Between Human, Technology and Organization in Building Information Modelling (BIM) - A Scoping Review of Critical Factors for the Individual User," *International Journal of Industrial Ergonomics* 97 (2023): 103480.

57. M. J. Page, D. Moher, P. M. Bossuyt, et al., "PRISMA 2020 Explanation and Elaboration: Updated Guidance and Exemplars for Reporting Systematic Reviews," *BMJ (Clinical research ed.)* 372 (2021): n160.

58. V. Braun and V. Clarke, "Conceptual and Design Thinking for Thematic Analysis," *Qualitative Psychology* 9, no. 1 (2022): 3–26.

59. D. Moher, A. Liberati, J. Tetzlaff, and D. G. Altman, "Preferred Reporting Items for Systematic Reviews and Meta-Analyzes: the PRISMA Statement," *International Journal of Surgery (London, England)* 8, no. 5 (2010): 336–341.

60. M. J. Page, J. E. McKenzie, P. M. Bossuyt, et al., "The PRISMA 2020 Statement: an Updated Guideline for Reporting Systematic Reviews," *BMJ (Clinical Research ed.)* 372 (2021): n71.

61. M. J. Page, J. E. McKenzie, P. M. Bossuyt, et al., "The PRISMA 2020 Statement: an Updated Guideline for Reporting Systematic Reviews," *Systematic Reviews* 10, no. 1 (2021): 89.

62. I. e. a. Goodfellow, "Generative Adversarial Nets," in *Proceedings of the International Conference on Neural Information Processing Systems* (Association for Computing Machinery, 2014), 2672–2680.

63. S. Colabianchi, A. Tedeschi, and F. Costantino, "Human-Technology Integration With Industrial Conversational Agents: A Conceptual Architecture and a Taxonomy for Manufacturing," *Journal of Industrial Information Integration* 35 (2023): 100510.

64. S. S. Sohail, F. Farhat, Y. Himeur, et al., "Decoding ChatGPT: A Taxonomy of Existing Research, Current Challenges, and Possible Future Directions," *Journal of King Saud University, Computer and Information Sciences* 35, no. 8 (2023): 101675.

65. V. Braun and V. Clarke, "Thematic Analysis," in *APA Handbook of Research Methods in Psychology, Vol 2: Research Designs: Quantitative, Qualitative, Neuropsychological, and Biological* (American Psychological Association, 2012), 57–71.

66. V. Braun, V. Clarke, N. Hayfield, L. Davey, and E. Jenkinson, *Doing Reflexive Thematic Analysis, in Supporting Research in Counselling and Psychotherapy* (Springer International Publishing, 2022), 19–38.

67. M. Naeem, W. Ozuem, K. Howell, and S. Ranfagni, "A Step-by-Step Process of Thematic Analysis to Develop a Conceptual Model in Qualitative Research," *International Journal of Qualitative Methods* 22 (2023), https://doi.org/10.1177/1609406923120578.

68. K. D. Dhole, "KAUCUS: Knowledge Augmented User Simulators for Training Language Model Assistants," in *Proceedings of the 1st Workshop on Simulating Conversational Intelligence in Chat (SCI-CHAT)* (Association for Computational Linguistics, 2024), 53–65, https://aclanthology.org/2024.scichat-1.5/.

69. M. A. Camilleri, "Factors Affecting Performance Expectancy and Intentions to Use ChatGPT: Using SmartPLS to Advance an Information Technology Acceptance Framework," *Technological Forecasting and Social Change* 201 (2024): 123247.

70. W. Elhag, D. Bouteldja, and S. Bouallegue, "A Survey on the Applications, Limitations, and Ethical Considerations of ChatGPT in Various Industries," in *2023 IEEE International Conference on Metrology for eXtended Reality, Artificial Intelligence and Neural Engineering (MetroXRAINE)* (2023), 565–569, https://doi.org/10.1109/MetroXRAINE58569.2023.10405608.

71. Z. Chen, Z. Jiang, F. Yang, et al., "Graph Meets LLM: A Novel Approach to Collaborative Filtering for Robust Conversational Understanding, 2023. [Online] https://www.amazon.science/publications/graph-meets-llm-a-novel-approach-to-collaborative-filtering-for-robust-conversational-understanding

72. N. Rane, S. Choudhary, and J. Rane, "Intelligent Manufacturing Through Generative Artificial Intelligence, Such as ChatGPT or Bard," *SSRN Electronic Journal* (2024), https://doi.org/10.2139/ssrn.4681747.

73. P. Maddigan and T. Susnjak, "Chat2VIS: Generating Data Visualizations via Natural Language Using ChatGPT, Codex and GPT-3 Large Language Models," *IEEE Access* 11 (2023): 45181–45193.

74. Y. Tang, A. A. B. Da Costa, X. Zhang, I. Patrick, S. Khastgir, and P. Jennings, "Domain Knowledge Distillation From Large Language Model: An Empirical Study in the Autonomous Driving Domain," in *2023 IEEE 26th International Conference on Intelligent Transportation Systems (ITSC)* (2023), 3893–3900, https://doi.org/10.1109/ITSC57777.2023.10422308.

75. E. Frydenlund, J. Martínez, J. J. Padilla, K. Palacio, and D. Shuttleworth, "Modeler in a Box: How Can Large Language Models Aid in the Simulation Modeling Process?," *Simulation* 100, no. 7 (2024): 727–749.

76. A. Kumar, M. L. Devi, and J. S. Saltz, "Bridging the Gap in AI-Driven Workflows: The Case for Domain-Specific Generative Bots," in *2023 IEEE International Conference on Big Data (BigData)* (2023), 2421–2430, https://doi.org/10.1109/BigData59044.2023.10386894.

77. E. Billing, J. Rosén, and M. Lamb, *Language Models for Human-Robot Interaction* (ACM, 2023), 905–906.

78. R. M. García Alarcia, P. Russo, A. Renga, and A. Golkar, *Bringing Systems Engineering Models to Large Language Models: An Integration of OPM With an LLM for Design Assistants* (SCITEPRESS—Science and Technology Publications, 2024), 334–345.

79. A. Shekhar, "Generative AI in Supply Chain Management," *International Journal on Recent and Innovation Trends in Computing and Communication*. 11, no. 9 (2023): 4179–4185, https://doi.org/10.17762/ijritcc.v11i9.9786.

80. Y. Xiang, Y. Wu, J. Song, Y. Gong, and P. Lianga, "Generative AI in Industrial Revolution: A Comprehensive Research on Transformations, Challenges, and Future Directions," *Journal of Knowledge Learning and Science Technology*. 3, no. 2 (2024): 11–20. https://doi.org/10.60087/jklst.vol.3n2.p20.

81. H. Aguinis, J. R. Beltran, and A. Cope, "How to Use Generative AI as a Human Resource Management Assistant," *Organizational Dynamics* 53, no. 1 (2024): 101029.

82. K.-B. Ooi, G. W. H. Tan, M. Al-Emran, et al., "The Potential of Generative Artificial Intelligence Across Disciplines: Perspectives and Future Directions," *Journal of Computer Information Systems* 65, no. 1 (2023): 76–107. https://doi.org/10.1080/08874417.2023.2261010.

83. D. Pradhan, B. Dash, P. Sharma, and S. Ullah, The Impact of Generative AI on Product Management in SMEs (2023). [Preprint] ResearchGate. [Online] https://www.researchgate.net/publication/378103355.

84. D. Van Hieu, "Incorporating Generative AI Into Quality Management Systems Enhancing Process Optimization and Product Development," *International Journal of Applied Machine Learning and Computational Intelligence* 13, no. 11 (2023): 1–8.

85. T. Osmeni and M. Ali, "Contemporary Generation: Artificial Intelligence Contribution to Manufacturing," in *2022 International Conference on Computing, Networking, Telecommunications & Engineering Sciences Applications (CoNTESA)* (2022), 31–35, https://doi.org/10.1109/CoNTESA57046.2022.10011233.

86. W. Wan, B. Liao, J. Cao, B. Zhang, L. Zhang, and J. Chen, "Transforming Automated Customer Service Responses Using ChatGLM Model: A Case Study of Expressway Toll Station Information System," *IEEE* 7 (2023): 1968–1973.

87. I. López-Riobóo-Botana, D. Gallent-Iglesias, and S. González-Vázquez, "QUA4I: Question Answering for the Industry 4.0 Domain. An Application of Intelligent Virtual Assistants," 2023.

88. V. Soni, "Impact of Generative AI on Small and Medium Enterprises' Revenue Growth: The Moderating Role of Human, Technological, and Market Factors," *Reviews of Contemporary Business Analytics* 6, no. 1 (2023): 133–153.

89. T. Süße, M. Kobert, S. Grapenthin, and B. F. Voigt, "AI-Powered Chatbots and the Transformation of Work: Findings From a Case Study in Software Development and Software Engineering," in *Collaborative Networks in Digitalization and Society 5.0, P. 2. Collaborative Networks in Digitalization and Society 5.0. 24th IFIP WG 5.5 Working Conference on Virtual Enterprises*, vol. 688 (Springer Nature Switzerland, 2023), 27–29.

90. S. Huo, K. Mukherjee, J. Bandlamudi, V. Isahagian, V. Muthusamy, and Y. Rizk, *Accelerating the Support of Conversational Interfaces for RPAs Through APIs*, vol. 491 (Springer Nature Switzerland, 2023), 165–180.

91. Q. Xu, G. Zhou, C. Zhang, F. Chang, Y. Cao, and D. Zhao, "Generative AI and Digital Twin Integrated Intelligent Process Planning: A Conceptual Framework," 2023, 29.

92. E. Brynjolfsson and L. R. Raymond, "Generative AI at Work (Working Paper)," NBER Working Paper Series, 2023.

93. K. U. Ahn, D. W. Kim, H. M. Cho, and C. U. Chae, "Alternative Approaches to HVAC Control of Chat Generative Pre-Trained Transformer (ChatGPT) for Autonomous Building System Operations," *Buildings* 13, no. 11 (2023): 2680.

94. C. Li, D. Chrysostomou, X. Zhang, and H. Yang, "IRWoZ: Constructing an Industrial Robot Wizard-of-OZ Dialoguing Dataset," *IEEE Access* 11 (2023): 28236–28251.

95. S. A. Prieto, E. T. Mengiste, and B. García de Soto, "Investigating the Use of ChatGPT for the Scheduling of Construction Projects," *Buildings* 13, no. 4 (2023): 857.

96. T. Medeiros, M. Medeiros, M. Azevedo, M. Silva, I. Silva, and D. G. Costa, "Analysis of Language-Model-Powered Chatbots for Query Resolution in PDF-Based Automotive Manuals," *Vehicles* 5, no. 4 (2023): 1384–1399.

97. C. B. Frey and M. Osborne, *Generative AI and the Future of Work: A Reappraisal (Working Paper)* (2023). Oxford Martin School. [Online] https://oms-www.files.svdcdn.com/production/downloads/academic/2023-FoW-Working-Paper-Generative-AI-and-the-Future-of-Work-A-Reappraisal-combined.pdf

98. Z. Wang, "Empowering Few-Shot Recommender Systems With Large Language Models-Enhanced Representations," *IEEE Access* 12 (2024): 29144–29153.

99. Y. Ye, H. You, and J. du, "Improved Trust in Human-Robot Collaboration With ChatGPT," *IEEE Access* 11 (2023): 55748–55754.

100. J. Yang, Y. Wang, X. Wang, X. Wang, X. Wang, and F. Y. Wang, "Generative AI Empowering Parallel Manufacturing: Building a "6S" Collaborative Production Ecology for Manufacturing 5.0," *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 54, no. 11 (2024): 6522–6536.

101. S. Fosso Wamba, M. M. Queiroz, C. J. Chiappetta Jabbour, and C. (V.) Shi, "Are Both Generative AI and ChatGPT Game Changers for 21st-Century Operations and Supply Chain Excellence?," *International Journal of Production Economics* 265 (2023): 109015.

102. L. Elhniy, X. Ye, U. Speidel, and S. Manoharan, "A Systematic Review of Recent Research on Chatbots," in *2023 International Conference on Electrical, Communication and Computer Engineering (ICECCE)* (2023), 1–9.