POLITECNICO DI TORINO

Master's degree in engineering and management

Master Thesis

Quality Engineering for GenAI-ALM Integration in Automotive Requirements Management

Academic Year: 2024 - 2025

Academic Advisor: Luca Mastrogiacomo

Business Advisor: Angelo Borneo

Candidate: Adele Fichera

Abstract

This thesis proposes a theoretical framework and a candidate architecture for the integration of Generative Artificial Intelligence (GenAI) within Application Lifecycle Management (ALM) platforms, to address the challenges, safety constraints and regulatory demands of the Requirements Management in the automotive industry. The research was conducted in collaboration with *MCA Engineering S.r.l.*, an international engineering and high-tech consultancy company based in Turin, operating across numerous innovation-driven projects in the automotive sector. The preliminary phase comprises a comprehensive literature review and market analysis, followed by the identification of the product quality model. The analysis prioritizes four of the nine characteristics defined by the ISO/IEC 25010 standard, namely: maintainability, reliability, security, and usability. In order to quantify each attribute, Key Quality Indicators (KQIs) were derived according to the mathematical functions outlined in the ISO/IEC 25023 standards. Adopting a design science research methodology, the project led to the formulation of a conceptual proposal which includes a make-or-buy assessment, a project roadmap, and a risk evaluation through Failure Modes and Effects Analysis (FMEA). The conceptual framework and the logical architecture contributed, arise in response to the immaturity of the current market landscape and the evident academic gap in the GenAI-ALM integration in the automotive sector. The model illustrates the system's core components, benefits, cost structure, and potential failure modes and outlines the key GenAI-powered functionalities such as requirement disambiguation, summarization of requirement documents, and translation of natural language inputs into technical specifications. The proposed solution will undergo a structured validation phase during the following internal deployment in MCA Engineering, through expert evaluation, to improve the system for future commercialization.

Summary

List of tables

List of figures

Introduction

In recent years, the Generative Artificial Intelligence technology is reshaping industrial priorities and operational dynamics. Among all sectors, the automotive industry stands out as one of the leading investors in GenAI, due to the increasing complexity across all its departments, from the early-stage conceptualization, through the production process, to the commercial field. Nowadays, modern vehicles integrate numerous systems, electronic control units, and software components, all of which must respect several constraints, concerning safety, functionality and data privacy. As a result, the Requirements Management methodology, which documents, traces, analyses and prioritizes requirements through the entire lifecycle development process, is one of the most affected fields by the increasing complexity in the automotive industry. The top challenges are the frequent changes, due to the rapid evolution of innovation, regulatory compliance and stakeholder misalignment. To prevent delays, cost overruns and compliance failures, Requirements Management tools based on the Application Lifecycle Management (ALM) are leveraged. These tools can ensure complete traceability across all requirements, facilitating the management of revisions and enhance global collaboration among distributed teams, improving operational efficiency through real-time integration capabilities.

Although, the application of GenAI into most of the automotive sector is increasing, its integration into ALM platform remains an emerging approach and does not yet constitute a mature, market-ready solution for Requirements Management. Current ALM platforms still rely on manual and time-consuming processes, often lacking automation, standardization, and intelligent support for complex engineering workflows. Moreover, there is a lack of academic models and industrial benchmarks addressing the quality, safety, and reliability implications of GenAI adoption in regulated domains such as the automotive one.

This thesis aims to propose a quality-driven framework and a candidate architecture for the integration of Generative Artificial Intelligence into ALM platforms, focusing on Requirements Management in the automotive context. The project was developed in collaboration with MCA Engineering S.r.l, an international engineering and high-tech consultancy company based in Turin, operating across numerous innovation-driven projects in the automotive sector. The objectives of this thesis include thinning the academic gap in this field and providing a project roadmap that can guide MCA Engineering S.r.l. into the integration of GenAI into ALM platform, fostered by a qualitative and risk analysis.

The research comprises firstly a literature review and market analysis to determine the actual conditions of the GenAI application in the overall businesses processes and in the automotive sector. Then, focusing on the GenAI integration within ALM platform to enhance Requirements Management, the ISO/IEC 25010 standard was the foundation to identify the four main features that this product should encounter, namely: maintainability, reliability, usability and security. After that, Key Quality Indicators (KQIs) were detected through the mathematical functions defined by the ISO/IEC 25023, which allow the quantification of the previous features and their corresponding sub-features. A Design Research approach was used to develop the proposal, which includes a four-phase project roadmap aligned with quality gate criteria, from requirements analysis to industrialization; a candidate architecture of the tool; a make-or-buy evaluation and a risk assessment through Cost of Quality identification and Failure Mode and Effects Analysis (FMEA). The latter includes the identification of potential failure modes and the proposal of corrective actions, which will be progressively implemented and verified across the subsequent phases of the roadmap.

This project derives from the strategic necessity of MCA Engineering to overcome the current challenges in Requirements Management faced during project development. The work proposed in this thesis represents the foundational step toward the implementation of an innovative, quality-oriented product designed to improve company's internal project management capabilities while, in the long term, enabling a competitive advantage through its future commercialization.

In the first chapter, this thesis introduces the theoretical background of Generative AI and its applications in industry. Then, presents an overview of the technologies, frameworks and providers, focusing on the types of GenAI, their foundation and adoption methods. The third chapter depicts the Automotive Sectors and its GenAI adoption, focalizing on the Requirement Management and the current situation. The alignment with the ISO/IEC 25010 and ISO/IEC 25023 is the centre of the fourth section, concerning the quality attributes and relative KQIs identification. The fifth chapter describes the State of Arts of GenAI-ALM integration in Requirements Management in the automotive sector, followed by the candidate architecture and the project roadmap. Lastly, the evaluation of benefits, risks and quality is developed, through a Cost of Quality (CoQ) identification and FMEA methodology.

1. Analysis of the GenAI Market for Business Use

1.1 What is Generative AI

Generative Artificial Intelligence (GenAI) is a branch of Artificial Intelligence focused on creating new content, such as text, images, code, or audio, through machine learning models trained on large datasets. GenAI use artificial neural network to identify the patterns and structure within existing data to generate new and original content. The artificial neural network (ANN) is a biologically inspired computational model that is modelled after the network of neurons present in the human brain, seen as a model of a parallel computational device. It transforms input data by applying nonlinear functions, *neural unit,* to their weighted sums. Each transformation is called *neural layer,* and the output of each layer is used as the input into the next one. During the training phase, the network modifies its weights to minimize the prediction error, enabling it to learn complex patterns. We can distinguish *feed forward networks*, which propagate information in one direction and *Recurrent Neural Networks (RNNs)* that include feedback loops, giving them memory to better handle sequential data [1].

The GenAI ability to create new, original content makes it a powerful tool in areas like entertainment, marketing, and design. Moreover, its capacity to simulate complex data patterns, allows researchers and businesses to uncovers otherwise hidden insights. Hence, GenAI is enhancing creativity, accelerating innovation and decision-making, making its application in industry a central focal point for all type of firms.

Although GenAI has made significant progress, is facing several challenges, such as the need for large-scale of compute infrastructure, significant capital investment, technical expertise, to maintain and develop generative models and massive compute power to train such large datasets. The lack of high-quality training data and the struggles of the organizations to get a commercial license to use existing datasets or to build datasets are critical bottlenecks limiting the GenAI training [1].

The existing body of research has recognized, among the different GenAI models, two prominent ones: the Generative Adversarial Network (GAN) and Generative Pre-trained

Transformer (GPT). A GAN employs two competing neural models to create new data similar to the original and fake data based on authentic data. GPT instead is s a pre-trained generative model based on multilayer deep neural networks [1]. It is designed by training extensive text datasets to understand language and generate new texts based on context and it relies on a single transformer model. The transformer model is a type of neural network architecture that excels at process sequential input data non-sequentially. It is used in numerous sectors like healthcare, bioinformatic, and finance, detecting trends and anomalies to prevent fraud [2].

The most popular generative AI applications [3]:

- Language: considered the most advanced domain, the text is at the root of many generative AI models. Large Language Model (LLM), represents the class of the transformer networks and is one of the most used to recognize, summarize, translate, predict, generate text and other forms of content-based knowledge gained from massive datasets.
- Audio: models able to develop song and snippets of audio clips with text inputs, recognize objects in videos and creates custom music.
- Synthetic data: useful to train AI models when data does not exist or are limited. The synthetic data is artificially generated to accelerate AI model training across many domains, and its development is the most impactful solutions for overcoming the data challenges of many enterprises [4].

1.2 How AI is Reshaping Business Operations

According to the latest McKinsey Global Survey on AI [5]: "*organizations are starting to make organizational changes designed to generate future value from GenAI, and large compagnie are leading the way*." The costly mitigation of GenAI-related risk along with the hiring of new roles capable of manage its use, makes the large companies the drivers of this transformation, followed by the SMEs (Small and Medium-size Enterprises).

The growing ambition to adopt GenAI is evident across every organization. *Gartner [6]*, leading technology research and advisory company, has provided a guide to help IT leaders define their "*AI ambition*," spot AI opportunities, prepare the right cybersecurity and adopt the AI principles. Indeed, 78% of the organization are planning to increase the investments in AI budget to pursuit new value creation, as confirms the Deloitte (2025)[7].

Even *Accenture [8]*, a global consulting firm specialized in technology and strategy, provides a roadmap based on six priorities with the purpose of supporting organizations to integrate the use of GenAI, in a reliable and effective way. First, organization are encouraged to combine quick wins gained using off-the-shelf models and strategic customized, data-driven models, to effectively define and deliver business value. A people-first approach highlights the importance of investing not only in the technologies, but also in talent development, to address two challenges: AI creation and adoption. The focus is also on the customization of cloud-based data platform, built on trusted and reusable data products, and on the adoption of a "*robust green software development framework"* that incorporates energy efficiency and low-emission principles across the software lifecycle, supporting ESG goals. Finally, success depends on the cooperation with strategic partners such as, technology company, professional services firms and academic institution, and on the establishment of effective governance structure to regulate risk management and compliance. Despite this though, the *Deloitte* (2025)[9] survey shows that "the main barriers today are *regulatory uncertainty and difficulty in demonstrating short-term economic value".* Demonstrating that, concerns related to compliance and risk management, increased respectively of the 10% and 6%, compared to early 2024. The difficulty in finding technical skills instead, decreased by 10 percentage points, indicating an improvement in the maturity of the sector [9]. The rapid advancement of GenAI technology, so much faster than organizational speed of the companies, represent another structural challenge which explains why many initiatives remain in the experimental phase and struggle to scale [7].

To respond to these concerns, the development of all the legal/ethical tools is increasing. Indeed, the 49% of the organizations have adopted AI ethics guidelines and the 37% are going to do so. Deloitte developed the "Trustworthy framework AI" which integrate ethical principles into the AI lifecycle, thanks to its six key dimensions:

- **Transparency**: reducing of "black boxing" and promoting responsible and traceable communication.
- **Reliability**: guarantee that AI technologies operate in a consistent manner, enhancing consumer trust.
- **Fairness**: reducing bias and discrimination, promoting AI models that respect the right of all individuals, treating them equitably.
- **Accountability**: define clear governance structure, roles, responsibilities and reporting mechanisms.
- **Privacy**: protecting the personal and sensitive data processed by AI systems, in compliance with current regulations.
- **Security**: ensuring the protection of AI systems and their infrastructure against cyber threats and unauthorized access. [9]

The increasing reliance on foreign technologies from dominant players like the U.S. and China and the relative growing concerns over data privacy, geopolitical tensions, create Sovereign AI movement [10]. This initiative aims to control the AI systems and ensuring that data, computing infrastructure, and technologies remain within the borders of nation. Sovereign AI is not just a necessity for regulatory compliance but also a strategic tool for nations to maintain economic competitiveness. Countries that embrace sovereign AI are likely to secure their place in the future of technology. In fact, states such as, Japan, Canada, the UK, and the EU are making large investments in AI infrastructure and local talent to ensure that they can develop AI solutions tailored to their needs while complying with local regulations. Hence, on one hand, there are huge, potential benefit of sovereign AI, such as the enhancement of control, security and compliance and the fostering of domestic innovation, thanks to the funds provided by the governments for research, development, and talent acquisition. On the other hand, developing and maintaining independent AI infrastructure while keeping pace with the rapid advances made by major global players such as U.S and China, is difficult and costly, not to mention the possible lack of interoperability among the different AI developed.

One of the main advantages of GenAI lies in its customizability. Each organization can fine-tune the "off-the-shelf" model by introducing their own data, allowing it to support specific tasks, and improving its responsiveness to signals of change. In fact, GenAI has the potential to reinvent its role in every company, it can carry out the mechanical tasks previously attributed to employees, leaving them with more managerial responsibilities, such as monitoring the accurate and responsible use of new systems based on artificial intelligence. Furthermore, advising, creating, coding, automating, and protecting are the key functions where the footprint of AI is the most visible.[8]

The most advanced GenAI initiatives, according to the *Deloitte 2025* [9], are:

- *Information Technology (28 %)*: leveraging GenAI to generate code and automate testing.
- *Operations (11%):* supports process optimization, predictive maintenance, and automation of repetitive tasks, in industrial sectors.
- *Marketing (10%):* enabling personalized content, predictive analytics, and campaign optimization at scale.
- *Customer Service (8 %):* thanks to the use of chatbots, autoresponder systems and tools for 24/7 support based on natural language.
- *Cybersecurity (8%):* used to handle large volumes of alerts, reduce false positives, and increase the effectiveness of defences [9].

1.3 Mapping of existing tools for business use

1.3.1 ChatGPT Enterprise

ChatGPT Enterprise was released in August 2023 by OpenAI, to meet the needs of business leaders who would like to deploy ChatGPT in their organization in a "simple and safe way". The model provides security and privacy, in fact, it is not trained on the business data of the organizations or on its conversation, it does not learn from its usage and it is certified SOC 2 (Service Organization Control 2) compliant, which ensures organizations manage customer data in accordance with five trust service principles: security, availability, processing integrity, confidentiality, and privacy. Among the many qualities of this tool, there are: the possibility to be customized and to fit in different types of organizations, besides size and sector, the ability to provide Advanced Data Analysis, browsing that are optimized for specific roles with a longer, more detailed prompts [11]. Using chat templates for internal collaboration is another advantage of the *ChatGPT Enterprise*, because it allows the users to manage the technology without needing advanced prompt engineering skills. Hence, thanks to *ChatGPT Enterprise*, businesses can improve their operations, enhance their efficiency and creativity, find new innovative content creation design and exploring answers to complex business questions. *Pwc* and *Moderna* are two examples of the many businesses that started a partnership with OpenAI already in 2024, before the release of ChatGPT-4o model, representing two reality that believed and invested in this technology [12].

Pwc, one of the world's leading professional services firms, operating across more than 150 countries and specializing in audit, tax, consulting, and advisory services, declared: "We are actively engaged in GenAI with 950 of our top 1,000 US consulting client accounts alongside discussing the use and implications of AI with many of our audit clients, emphasizing the near universal demand across industries for the transformative power of this technology." [13]. In this way, *Pwc* helps its customers to evolve with AI, transforming its internal structure, redefining internal processes, enhancing quality and velocity maintaining responsible AI approach [13].

Moderna is biotechnology company specialized in the development of mRNA-based therapeutics and vaccines. It is using GenAI, particularly OpenAI's technologies, to redesign all business processes, from legal to R&D and manufacturing. Their transformation strategy included training programs, cultural change management, and executive involvement. The internal chatbot *mChat,* built on OpenAI's API and adopted by over 80% of employees, is the results of this strategy. This tool paved the way for the subsequent deployment of *ChatGPT Enterprise* and of GPT pilot called *Dose ID* within *ChatGPT Enterprise* to support clinical trial development. The tool reviews and analyses clinical data, integrates large datasets, and generates customized visualizations to assist study teams. By leveraging advanced data analysis, *Dose ID* automates dose selection processes, applying standard clinical criteria while providing rationale, references, and visual outputs. This allows researchers to explore data from multiple perspectives and validate optimal vaccine doses before late-stage trials. *Dose ID* does not replace clinical judgment, instead enhances efficiency, accuracy, and safety. According to Moderna, GPT enables faster, more comprehensive evaluations of vast datasets while ensuring data security and privacy [14].

1.3.2 Breeze

Breeze is a GenAI tool, introduced at HubSpot's INBOUND 2024 conference, designed to augment sales, marketing, and customer service processes directly within CRM ecosystem. It's composed by three main pillars: *Breeze Copilot, Breeze Agents* and *Breeze Intelligence*. The first one is an AI assistant embedded across HubSpot applications that provides contextual support such as drafting emails, summarizing records, or preparing meeting notes, without the requirement of AI expertise. The second one is a specialized AI modules that automate domain-specific tasks like creating marketing content, managing social media posts through performance analysis, industry data and best practices or by assisting customer interactions. Lastly, *Breeze Intelligence* enriches CRM data through features such as buyer intent detection and contact enrichment. By combining automation, personalization, and data-driven insights, *Breeze* enables organizations to streamline workflows, improve lead conversion, and achieve measurable productivity gains while maintaining a seamless integration with existing HubSpot environments and the stories of *Sandler* and *Agicap* are the evidence [15].

Sandler, a global provider of sales training and professional development, adopted *Breeze Copilot* within HubSpot's Marketing and Sales Hubs, achieving a 25% increase in customer engagement and generating four times more qualified leads. *Sandler's* team has become a trusted advisor, correcting GenAI misunderstandings, allaying fears, and demonstrating its true potential. Another use of *Breeze* is the opportunity to provide customized experience to its clients and enhance the empathy along the sales path. The AI managed data-driven approach, able to respond to the concerns of the customers with concrete data and case study, allows the reduction of 50% of the sales cycle[15].

Agicap is a financial technology company which adopted HubSpot's *Breeze* to reach efficiently the mid-market clients. The company saved over 750 hours weekly and achieved 100% CRM adoption across six countries by automating routine tasks such as call summaries, email drafting, and workflow management. *Breeze* also boosted deal velocity by 20%, enabling sales teams to focus on high-value interactions and deliver more personalized customer experiences [15].

1.3.2 Timely

Timely is an operation and time tracking platform based on GenAI, provided in 2024 by the Norwegian technology company Memory AS, with the aim of introducing automatic, AI-powered time tracking to enhance productivity and workflow management. The GenAI generated timesheet, learn from the previous registration models, allowing users to create and validate new report in seconds. Managers can have the real-time visibility into team performance, resource allocation, and project progress, thanks to centralized dashboards. Moreover, the additional features such as capacity planning, task management, project dashboard and capacity planning, task management, project dashboards, billable rates optimization, and integrations with over 100 apps extend its utility to complex business environments. Privacy-first by design, *Timely* emphasizes ethical self-tracking, making it particularly relevant for consulting firms, creative agencies, and software development teams [16].

The use case of *Soar With Us* and *Zagaran* demonstrate the incredible useful work of *Timely*. The former, with automatic background tracking and customizable reports, saved time, restored trust in its data and fostered a cultural shift where time tracking is now seen as a valuable, non-intrusive practice. The latter instead, a Boston-based software consulting firm, by adopting *Timely* and its features like calendar integration, mobile logging, and API customization, improved data accuracy, resource allocation, and client billing. *Timely* has since become integral to their operations, enabling scalability and higher client satisfaction [16]

1.4 Trends and emerging use cases

The exceptionally fast development of AI technologies makes even the most recent considerations and research on trends and trajectories quickly obsolete. Research and applications are numerous and highly fragmented, so that what is already being applied in one company may still be at the research stage in others. The following sections describe the most common trends of this new technology.

1.4.1 Agentic AI

The *AI agent* exemplifies the phenomenon mentioned above: while in 2024 the global management consulting and professional services firm ZS, argues that current enterprise applications of GenAI remain confined to basic chatbots and content generation, firms such as Breeze are already delivering specialized agents and leveraging their impact for clients worldwide [17].

The AI agent is a fine-tuned GenAI equipped with different capabilities to enhance decision-making and perform specific tasks, relative to a specific *context* and environment. Among the others there are: *persona*, with guideline to plan in specific ways, *memory,* in which the GenAI agents retain memory of previous interactions allowing coherence and continuity, *knowledge* where can access structured data or vector databases and *tool integration*, which allows the agents to take actions using those tools.

As highlighted by the CIAIC (Collective Intelligent of AI Consultant) framework (2024), AI agents can collaborate as specialized consultants and deal with the real-world use cases which require coordinated efforts from multiple AI agents to achieve transformative outcomes. In multiagent workflows, collectively refining strategic outputs by gathering and integrating data in real-time, ensuring that decisions are based on the most relevant and up-to-date information [18]. Key areas of transformation include supply chain resilience, regulatory document creation, clinical trial optimization, marketing content creation, patient support, and software development [17]. Although, multiagent collaboration is poised to disrupt and enhance various enterprise processes, the CIAIC framework ensures human involvement at the initial stages of strategic planning, with experts defining the objectives and refining the outputs, making the process collaborative, and ensuring the final strategic plans are comprehensive and reliable [18].

1.4.2 Physical AI

Physical AI integrates artificial intelligence with robotics, autonomous vehicles, the Internet of Things (IoT), and digital twins to sense, interpret, and act in the physical world. Manufacturing, logistics and health care are the three main industries which have been transforming by physical AI, reducing defect and downtime, using drones for deliveries and warehouse operation, and enhancing health care responsiveness, through wearable sensor able to monitor the patient in real-time [19]. According to the International Federation of Robotics (IFR), adopting these machines could boost productivity by 20-30% in key industries by 2030. The U.S. and Chinese innovators are focusing on the $30 trillion global labour market, deploying robots equipped with machine learning, computer vision, and advanced sensory systems. However, this rise in efficiency and productivity brings its disadvantages. Repetitive manual jobs, for instance, are expected to shrink by 10-15% globally within a decade, as stated by the IFR. This shift will likely put downward pressure on wages in low-skill roles such as packaging and assembly. Simultaneously, demand for technicians, software developers, and engineers to manage and oversee these systems will increase significantly. These roles will require specialized skills, which displaced workers may struggle to acquire without robust support. Moreover, privacy and data security remain pressing concerns. Companies must collaborate with regulators to establish clear laws, ethical frameworks, and safety standards, particularly in relation to humanoid robots used in both domestic and workplace environments [20].

1.4.3 Multimodal AI

Multimodal AI models can interpretate and process different types of information generating outputs based on heterogeneous inputs. They are composed by multiple neural networks, each specialized in a specific format (test, audio, images, video, …). In the initial phase, data of different format are pre-processed and then encoder tools transfer them into machine-readable feature vectors or *embeddings* by the encoders within each network. Subsequently, the encoded data are projected into a shared space using various fusion mechanisms, which merge the embedded text from different modalities into a layer, allowing the model to understand the relationship between the different data and focus on the most relevant. The last phase, the generative one, converts the fused data into actionable outputs, using autoregressive methods or generative adversarial networks (GANs).

The adoption of multimodal GenAI offers several strategic applications for organizations, accelerating creativity in marketing and product design, enabling personalized campaigns and rapid prototyping. In the insurance sector, it can mitigate fraud by monitoring diverse data sources like statements, and images. It also enhances trend detection by analysing unstructured data from social media and multimedia content. In healthcare, multimodal GenAI has the potential to transform patient care with intuitive multimodal virtual assistants and real-time call centre support. Moreover, it improves accessibility and user interaction by enabling automated assessments among applications and platforms. Companies that invest early in these applications may gain a significant competitive advantage. The risks associated with multimodal AI are like those of other forms of GenAI but often more severe. Hallucinations, for example, can lead to cascading errors across complex systems, while handling sensitive multimodal data heightens privacy and security concerns. Bias and fairness remain critical issues, as integrating heterogeneous data demands careful system and model architecture design. Furthermore, evolving regulations on AI and data use add an additional layer of complexity across industries and regions. To mitigate these risks, organizations should adopt reliable and updated models, implement robust guardrails, maintain human oversight in sensitive tasks, and focus on low risk use cases where errors are easily verifiable or have minimal impact [21].

2 Classification of Technologies, Frameworks, and Providers

2.1 Foundation and enterprise pretrained model.

In recent years two fundamental concepts have been at the centre of the debate on the GenAI development: foundation and pre-training models.

Foundation models are a large-scale neural network architecture, trained on enormous raw data through self-supervised learning, adapted to accomplish a broad range of tasks. The primary purpose of foundation models is to act as common basis from which many task-specific models are built via adaptation [22] . It's a perfect building block for wide range of applications, including question answering, text generation, summarization, and even code generation. Self-supervised learning allows these models to acquire meaningful representations from the input data, without relying on explicit labels or annotations [23]. During the pretraining phase, the model enhances patterns and relationships within the data, thanks to many processes such as predicting missing parts, fixing corrupted data, or finding connections between different data types. For instance, a model pretrained on textual data might understand word relationships, sentence formations, and diverse writing styles. This foundational understanding enables the model to perform tasks like generating text, translating content, and answering questions after fine-tuning. Similarly, a model trained on diverse image datasets, can learn to recognize objects, their parts, and their underlying relationships [22].

Other three key characteristics of AI foundation models include [23]:

- Fine-tuning: process that enables the model to adapt their pre-existing knowledge to specific tasks and to improve its performance on definite objectives.
- Multimodal capabilities: the model is designed to handle multiple modalities such as text, images and video, simultaneously.
- Transfer learning: which is the ability of applying knowledge learned from one task to another related one, improving efficiency and performance on new but similar challenges.

Despite their transformative potential, foundational models also raise significant complications. One significant risk is the amplification of biases embedded in the large datasets used for training, together with the introduction of inaccurate or misleading information in images or videos and the violation of intellectual property rights of existing works. As emphasized in the Standford report on foundation models: "*future AI systems will likely rely heavily on foundation models, it is imperative that we, as a community, come together to develop more rigorous principles for foundation models and guidance for their responsible development and deployment*" [24].

Pretrained AI models are machine learning architectures trained on large-scale datasets to accomplish a *specific task*. Developers can rely on these models that have already learned general features, such as language structure or visual shapes, and fine-tune them on smaller, domain-specific datasets. This approach accelerates development and allows smaller entities, such as startups without access to sufficient computers, data or infrastructure, to experiment with state-of-the-art systems [25]. Transfer learning applies this idea by using the knowledge learned from pre-trained models to solve new, related tasks, even when only a limited amount of additional labelled data is available. Since pre-trained models are usually trained on large datasets curated by domain experts, developers can also leverage this expertise without having to create their own datasets. This makes pre-trained models useful in specialized domains where labelled data is scarce and creating labelled datasets is difficult. As such, pre-trained models can act as a foundation for further training to develop innovative AI applications in a wide range of industries [26].

One of the first pretrained models and open-source AI system is BERT (Bidirectional Encoder Representations from Transformers), released by Google in 2018 [27]. State-of-art language model, it is designed to pre-train deep bidirectional representations using both left and right context in all layers; thanks to that, achieves superior performance on several NLP (Natural Language Processing) tasks, including question answering, natural language inference, and sentence classification [27].

In organisation contexts, large foundational models are increasingly embedded within enterprise platforms (e.g., Salesforce, Microsoft, SAP), where they are combined with proprietary customer data to produce industry-specific applications. This practice, often referred to as *enterprise pretraining*, enables companies to adapt general-purpose models to their unique business needs, thereby unlocking substantial value. However, it also raises critical challenges, such as risks of exposing sensitive or proprietary data, compliance with regulations like GDPR (General Data Protection Regulation), and the amplification of bias or hallucinated content. To address these concerns, platform owners are introducing new architectural layers, including trusted environments to guarantee secure data access, privacy, and content moderation, as well as prompt architectures that improve accuracy and relevance of outputs. Moreover, effective adoption of pretrained models in enterprise settings requires dedicated governance mechanisms to regulate data access, ensure transparency of AI-generated content, and promote the responsible and sustainable use of generative AI technologies [28].

2.2 Open-source vs commercial AI

Commercial AI solutions and Open-source LLMs represent the two most adopted GenAI forms among businesses. Both are primarily based on foundation or pretrained models but differ in terms of accessibility, support and additional functionalities. Open-source models, such as GPT-2 and LLaMA, provide flexible and customizable foundation, while the commercial ones, like GPT-3, are optimized for business applications and come with dedicated support infrastructure. Choosing between proprietary and open-source tools it’s difficult and depends on the business’s priorities, objectives, and available resources.

Commercial GenAIs are provided from private entities, featuring advanced support and specialized service, usually accessed via APIs (Application Programming Interface) or licensed software, managed and maintained by a company [29]. They allow the organization to integrate them without significant investment in time or resources. Many of these services, such as ChatGPT itself, have high-standard security controls and adhere to data regulations. The main risk of these systems is the possibility that sensitive information could be used in training datasets; besides the fact that they can generate inaccurate responses since they are not fully aligning with the specific enterprise policies [30].

A key example of commercial AI integration is the collaboration between Mercedes-Benz and OpenAI. In this case, ChatGPT is integrated into its MBUX infotainment system to enhance the driver experience by providing a conversational interface and allowing for hands-free control of various functions such as navigation and climate control. Multimodal integration, personalized experience, integration with Other Service and enhanced voice assistant are the key features of ChatGPT in MBUX which leads to its advantages, such as: the ease of use, where system understands context and can provide more natural interactions based on the driver's preferences; the enhancement of safety thanks to the reduction of the need for manual inputs, minimizing distractions; the continuous learning because over time, the AI assistant becomes more attuned to user preferences and driving patterns, offering increasingly accurate responses and tailored recommendations [31]. Even Audi is enhancing its infotainment system by integrating ChatGPT via Microsoft Azure OpenAI Service. As seen in the previous example, safety, efficiency, and enhanced voice interaction are the main benefits of this integration. Additionally, the system ensures data privacy by deleting all interactions with ChatGPT after processing, with the tool never accessing vehicle-specific data, ensuring a secure experience for users. The integration seamlessly works with the Audi assistant, which intelligently determines when to execute vehicle functions (such as adjusting the AC or setting navigation) and when to consult ChatGPT for general knowledge inquiries[32].

Open-source LLMs provide significant scope for implementing robust security measures and data protection protocols tailored to an organization’s requirements. They allow businesses to tailor GenAI models to their specific needs, enhancing the accuracy and relevance of the outputs. Meta’s LLaMA 2 is one of the most popular examples of Open-Source, designed for human-like text generation and dialogue, without compromising safety and utility. However, fine-tuning these models is intricate and resource-intensive, with a risk of producing biased or inaccurate results if not managed correctly. Choosing between these technologies is about strategically aligning AI capabilities with the business goals and risk profiles of each business. Among data privacy, cost implementations, and desired level of control over AI systems, scale and task specificity, are the two that the most guide the choice between open source and commercial options. The scalability is a high advantage of the Open-source AI models, that are adept at parallelizing thousands of requests, a necessity for large-scale operations; in contrast to commercial AI models which typically impose rate limits. Open-source AI is also preferable for tasks requiring specialized knowledge or expertise, they can precisely tailor and trained on specialized datasets, unlike the proprietary system that are designed for broad applicability but not for the niche requirements with the same level of specificity [29].

Among the most notable open-source GenAI technologies are BLOOM and Falcon 180B. BLOOM is an autoregressive Large Language Model (LLM) capable of generating coherent text across 46 languages and 13 programming languages. Trained with vast amounts of textual data using industrial-scale computational resources, BLOOM can continue text from a prompt and can also perform tasks it has not been explicitly trained for by framing them as text generation tasks. Its applications, beyond text generation, include language exploration, and a variety of downstream tasks such as information extraction, question answering, and summarization [33].

Falcon 180B, developed by the Technology Innovation Institute (TII), is another powerful open-access LLM. Falcon 180B can generate human-like text and power advanced natural language applications. Its pretraining data primarily consists of web data, supplemented with curated conversational and technical datasets, and a small fraction of code [34]. While Falcon 180B can be used commercially, usage is restricted, and hosting of the model is limited by license terms.

2.3 Costs and adoption methods

Nowadays, business leaders face a fundamental choice: leverage open-source models in-house or move towards commercial AI APIs. Some companies lean on the latter offering for advanced features and easier scaling, others prefer open source for its customizability and lower marginal cost. The proprietary model, on the other hand, ensures quick development, enterprise-grade support and cutting-edge updates.

Open-source AI can be an excellent choice for companies that handle large volumes of data or require stringent data governance. By hosting models in their own environment, avoid per-request fees and maintain full control over their data pipelines and infrastructure. While commercial APIs charge per call or per token, open-source models incur mostly fixed infrastructure costs, as GPU clusters or cloud with no surprise overage charges. Additionally, opens-source tools can leverage economies of scale becoming up to the 50% cheaper than an equivalent commercial API service [35].

In fact, research from the Linux Foundation highlights that open-source AI significantly reduces costs for companies [35]. The 60 % of respondents in McKinsey's survey confirmed that open-source AI has lower implementation and maintenance costs compared to the commercial options. Notably, open-source AI shows a slight edge in cost savings, with a typical improvement of 4% over proprietary tools. Marketing departments have experienced up to 5% higher savings with open-source solutions, while service operations have seen 6% higher revenue growth compared to proprietary systems [36].

2.3.1 Commercial AI

The costs for using commercial AI tools vary based on the type of content generated, text, audio, image, the provider and on which model. Moreover, at an organizational and resource level, using the basic version is inconvenient, so, it would be necessary to add a fine-tuning process with the purpose of merging the tool with data and business needs.

Considering first a commercial GenAI, we see that most of them leverage the pay-as-you-go monetization strategy, which means that charges businesses based on the number of charters or tokens in input or output text. For companies, predict how many queries employees will run, is challenging, mostly if they explore GenAI use cases in various departments. Another disadvantage is the lack of contextual knowledge that characterized the "off-the-shelf" tools, such as company structure, products and services, that unable the company to leverage the GenAI support. According to Eric Lamarre, senior partner at McKinsey [37], to solve this problem, companies must retrain commercially GenAI tools on their corporate data to improve accuracy and reliability. There are two ways to do that:

1. Using software-as-a-service (SaaS) platforms with generative AI capabilities. Prominent SaaS vendors like SAP, TIBCO Spotfire, and Salesforce are introducing generative AI services that can be fine-tuned using customer data. Salesforce's Einstein Copilot, part of the Einstein 1 Platform, offers an enterprise-grade AI assistant for creating custom models and prompts. While pricing is not disclosed, earlier pilot programs were reported to cost around $500 per user per month, with current prices varying based on configuration and agreements [38].
2. Integrating corporate software with Gen AI solutions over APIs and retraining models on business's data. To reduce GenAI implementation costs, companies can integrate commercial AI solutions directly via APIs [38]. The synchronization of ChatGPT-4o is one of the main examples, where the fine-tuning cost is of 25$ for million tokens, inference is $3.75 per million input tokens and $15 per million output tokens [39].

2.3.2 Open-source AI

Training a foundation model from scratch can require millions of dollars, making it prohibitive for many businesses. However, companies can leverage existing open-source model as a starting point to solve complex problems, with chance for customization. After selecting an open-source model, businesses can integrate it with their software using APIs and utilize their own server infrastructure.

This approach includes the following costs [38]:

- Hardware costs: these depends on the size and complexity of the model. Smaller models, can be trained using a high-end GPU, such as an NVIDIA RTX 3080, priced between $700 and $1,500. For larger models, like GPT-2, multiple high-end GPUs or specialized AI accelerators are required. For example, a single NVIDIA A100 GPU costs from $10,000 to $20,000, while a full multi-GPU setup may range from $30,000 to $50,000.
- Cloud computing costs: renting a cloud computing resources from providers like Amazon Web Services (AMS) is an alternative to buying hardware. Cloud-based GPU instances can range from $3 to $24 per hour depending on the instance type.
- Electricity and maintenance.
- Integration and deployment.
- Data storage and management. For on-site installations, storing generative AI data could cost between $1,000 and $10,000, depending on the dataset's size and redundancy needs. Cloud storage services like AWS S3 charge between $0.021 and $0.023 per GB per month, with additional costs for data transfer and operations.

For a mid-sized enterprise aiming to deploy a moderately large model like GPT-2 on-premises, the estimated generative AI costs would range from $37,000 to $100,000 for initial deployment, which includes hardware, integration, and data storage setup. Recurring costs, including electricity, maintenance, and ongoing integration, would range from $7,000 to $20,000 annually.

Considering fine-tuning an open-source foundation model, the costs increase due to several key factors, such as[38]:

- Model size: since the cost of training LLM increases with the size and complexity of model.
- Computational resources, determined by whether you use your own hardware or cloud services. For simpler models like local GPT-2 may require GPU investments of $10,000–30,000, while cloud computing costs for high-end instances can range from $1 to $24 per hour, depending on the instance type. Fine-tuning larger models like GPT-3 requires advanced GPU setups that can exceed $50,000–100,000.
- Data preparation: process of collecting, cleaning, and preparing your data for fine-tuning foundation models can be resource intensive.
- Development time and expertise.
- Maintenance costs.

For a mid-sized enterprise looking to fine-tune a moderately large model like GPT-2, the implementation of generative AI involves a range of costs. Initially, the investment in hardware can range between $20,000 and $30,000, depending on the specifications of the GPU setup needed to support the model's computational requirements. The development phase, which assumes a six-month timeline, can involve both in-house and outsourced talent. For in-house teams, the cost is estimated between $35,000 and $100,000 for a half-year salary [38]. Alternatively, outsourcing this work can cost between $20,000 and $40,000, assuming 400 hours of work at an average rate of $75 per hour.

Data preparation can vary widely depending on the size and complexity of the dataset. This step is expected to cost between $5,000 and $20,000. Furthermore, the ongoing maintenance of the system adds an additional cost of $5,000 to $15,000 per year, covering updates and necessary adjustments to ensure the system runs smoothly.

Hence, the initial deployment of the generative AI solution could range from $80,000 to $190,000, factoring in hardware, development, and data preparation expenses.

Recurring costs, such as maintenance and other operational expenses, are expected to range from $5,000 to $15,000 annually [39].

While AI adoption can be a significant investment, the potential benefits in terms of efficiency, innovation, and competitive advantage often justify the expenditure. Companies must conduct a detailed cost-benefit analysis and develop a clear AI strategy to ensure a successful and cost-effective implementation. By leveraging cloud-based AI services, strategic partnerships, and automation tools, businesses can optimize their AI investments and maximize long-term value.

3. Automotive Sector and GenAI Integration

In the current era of artificial intelligence, the challenge of adopting accelerated innovation remains a constant imperative for businesses. Among the sectors most actively embracing technological integration, the automotive industry stands out.

Generative AI is rapidly emerging as a key enabler, revolutionizing the design, production, sales, and usage of vehicles. It has become indispensable to the development of software-defined vehicles (SDVs), advanced driver assistance systems (ADAS), autonomous driving (AD), and in-vehicle customer experiences. Proof of this is the recent study by McKinsey, involving automotive and manufacturing executives that revealed: “more than 40% of respondents are investing up to €5 million in GenAI research and development, and more than 10% dedicating over €20 million” [37]. Such significant investments reflect a strong belief in the transformative potential of GenAI, as businesses seek to acquire the necessary technologies to address challenges including budgetary constraints and time-to-market pressures, while simultaneously driving transformation and meeting customer expectations. This technological innovation assists automotive manufacturers in achieving significant time and cost savings, alongside improvements in product quality, by adding value across all phases of the research and development lifecycle.

3.1 Adoption and Implementation of GenAI in Automotive

Generative AI is reshaping the automotive industry, influencing both production and customer experience. In design and engineering, it enables the creation of innovative, efficient vehicles with optimized weight, materials, safety, and cost. In production and supply chain management, GenAI analyses large datasets to identify inefficiencies, optimize resource planning, forecast component demand, and reduce delays and waste. It also supports predictive maintenance by anticipating failures through telematics and sensor data, minimizing downtime. In advanced driver assistance systems and autonomous driving, GenAI enhances environmental perception and object recognition, while digital twins and virtual simulations allow thousands of road scenarios, including extreme cases, to be tested, accelerating development without relying solely on real-world trials [40], [41].

Customer experience benefits from intelligent voice assistants, predictive navigation, and adaptive infotainment systems that personalize interactions based on context and driver preferences. Software development is streamlined through automated code generation, testing, and documentation for embedded systems, ADAS, and battery management. Moreover, GenAI contributes to sustainability and material innovation, supporting more efficient and environmentally friendly vehicle designs and production processes. Overall, GenAI enables automotive companies to improve efficiency, safety, product quality, and customer satisfaction while gaining a competitive advantage [40].
The following sections will focus on three key applications, software development, infotainment and quality control in manufacturing; to illustrate how generative AI is transforming the automotive sector.

3.1.1 Software development

Generative artificial intelligence (GenAI) is increasingly influencing the automotive software development landscape by optimizing various stages and significantly reducing the time required by developers. This transformative technology impacts the entire software lifecycle, from requirements management to architecture design and implementation. GenAI, especially through Large Language Models (LLM), facilitates the automation of requirement extraction and summarization, thus reducing the analysis of extensive documentation, which often involves thousands of specifications per vehicle. By analysing large datasets, GenAI enables more accurate translation of business needs into technical specifications, minimizing communication issues. Especially, models such as GPT-4 and other advanced LLMs can process vast amounts of textual data and generating precise, relevant software requirements. This automation reduces significantly the time and effort typically allocated to manual documentation analysis, making the translation of business needs into specific executable software more efficient [42].

Code generation is arguably the primary application of Generative AI (GenAI) in software engineering. It has the potential to significantly enhance developer productivity by reducing the time spent on various software engineering tasks, such as generating initial code drafts, correcting or refactoring code, and creating new system designs. During the development phase, GenAI supports writing, translating, refactoring, and documenting code. It can accelerate the overall coding process by generating, completing and creating code from pseudocode prompts. For critical software, such as Advanced Driver-Assistance Systems (ADAS), GenAI can produce optimized code for systems with limited memory and processing power, while ensuring appropriate hardware interfaces. Additionally, GenAI aids in translating code between programming languages, facilitating the modernization of legacy systems and enabling developers to address backlog tasks without rewriting existing codebases. It also identifies and improves problematic areas of code, enhancing maintainability, readability, and reducing technical debt [42].

Volkswagen has made significant investments in Generative AI to enhance its software development process. The company is utilizing Large Language Models (LLMs) to improve Requirements Management and automate code generation for vehicles. The company aims to integrate GenAI into its software production processes to optimize development time and reduce costs, particularly in the design and manufacturing of vehicles. This includes automating tasks such as requirements extraction, documentation analysis, and code drafting to enhance productivity across the development lifecycle [43].

3.1.2 Infotainment

Vehicle infotainment is another significant sectors where generative AI has been applied. By analysing user preferences, GenAI enables a personalized experience, suggesting music, podcasts, and audiobooks tailored to the individual. Moreover, it can adjust infotainment options in real time based on factors such as mood or driving conditions, ensuring that the content remains relevant and engaging during the journey. Automakers, including Mercedes-Benz, integrate GenAI to enhance interactivity through touchscreens, voice control, gesture recognition, smartphone integration, cloud services, and other IoT devices. The primary goal of these advancements is not solely customer comfort, but also safety. GenAI provides drivers with essential information without requiring them to divert their attention from the road, thereby minimizing distractions [44].

As the previous example of Mercedes-Benz and Audi, in 2025, Volvo Cars will introduce Google Gemini into its vehicles, integrating advanced generative artificial intelligence into the in-car infotainment system. Google Gemini will enable drivers to perform tasks such as sending messages, translating text, or accessing the car's manual simply by speaking, with the objective of reducing cognitive load and improving safety. Additionally, Volvo will collaborate with Google to develop new features, marking an important step in the automotive sector regarding connectivity and GenAI adoption, ultimately creating a more intuitive and safer driving experience [45].

3.1.3 Manufacturing

In automotive manufacturing, GenAI can be applied in several aspects, including design, supply chain management, quality control and predictive maintenance. One of the main areas of use is assisted design, where generative models can be used to create new variants of automotive components, optimizing the shape and functionality of parts so that they meet performance, strength and cost requirements. GenAI can simulate and generate innovative designs that reduce weight and improve fuel efficiency, while maintaining component quality and safety. Additionally, these models can diminish the time required for the design and testing cycle, speeding up the entire development process. In the field of production and supply chain management, GenAI is used to predict component demand, optimizing resource planning and logistics. By using predictive models, companies can avoid delays in production time, ensuring that the necessary materials are always available without overloading inventory, thus reducing operating costs. GenAI also helps to respond quickly to unforeseen changes in demand or disruptions in the supply chain, improving overall production efficiency. In the predictive maintenance industry, GenAI helps continuously monitor equipment and machinery used during production, predicting possible failures before they happen. AI analyses massive amounts of data from machinery, such as vibration, temperature, and noise, to identify signs of stress or malfunctions. In this way, companies can plan maintenance interventions without interrupting production, reducing costs and improving the reliability of the entire production plant. Moreover, GenAI can contribute to advanced quality control. Generative AI systems can analyse data from sensors during production and identify defects or anomalies in components with greater precision than traditional methods. For example, it can be used to improve automatic visual inspection of produced parts, identifying microdefects that could escape human operators or traditional computer vision technologies. This not only reduces the error rate but also helps ensure that each vehicle meets the highest quality standards. Mass customization is another advantage that GenAI offers in automotive manufacturing. AI can be used to generate customized options that are efficiently integrated into the production line, improving the customer experience without compromising production efficiency [41].

A concrete example of how smart technologies can be integrated into a connected factory is the BMW Group's implementation of the "GenAI4Q" project at its Regensburg plant. This generative artificial intelligence system optimizes quality control in vehicle assembly, analysing a vast amount of data from each vehicle produced, including model, variant, equipment and real-time production data. Based on this analysis, the system generates customized inspection protocols for every vehicle, autonomously determining both the scope and sequence of required checks.

This approach reduces the risk of errors and speeds up the inspection process. The system's intuitive interface, accessible via mobile devices, allows operators to efficiently record inspection results and incorporates advanced features such as voice recognition and automatic transcription. By leveraging GenAI in this manner, BMW can maintain high-quality standards across highly flexible production lines, simultaneously assembling internal combustion, plug-in hybrid, and fully electric vehicles with numerous customized configurations [46].

3.2 Requirements Management in the Automotive Sector

Requirements Management refers to the systematic process through which development teams document, trace, analyse, prioritise, and align requirements through the entire product development lifecycle. This process constitutes a fundamental component of the product life cycle, spanning from initial conception, through design, to the creation of the final product [47].

In the automotive sector, the value of Requirements Management is further amplified by the complexity of modern vehicles, which integrate numerous systems, electronic control units, and software components, as well as by strict regulatory and safety obligations. Automotive Requirements Management ensures that safety, functionality, and compliance objectives are outlined, verified, and tracked throughout the product life cycle, supporting thorough impact analysis, variant management, and compliance reporting.

According to *Visure* [48], the main functions of the Automotive RM are:

- identify regulatory and stakeholder requirements
- creating software requirements from high-level targets
- managing evolutionary product changes between the various teams collaborating
- support for audits and compliance with critical security standards.

The growing complexity of vehicles makes automotive RM face several challenges. According to *EnCo Software GmbH [49]* the top challenges are the frequent changes, due to the rapid evolution of innovation, regulatory compliance, stakeholder misalignment, caused by the poor communication between teams, suppliers and customers during the development process and the management of the thousands of interconnected systems.

3.2.1 The Role of Requirement Management Tools

As discussed in the previous section, keeping up with the continuous evolution of innovation and the corresponding requirement within the automotive industry presents significant challenges. Hence, leveraging Requirements Management tools based on the Application Lifecycle Management (ALM), such *Enco's SOX* [49]and the *Visure Requirements ALM Platform* [48], has become essential to prevent delays, cost overruns, and compliance failures.

A key advantage of these tools is their ability to ensure complete traceability across all requirements, facilitating the management of revisions and ensuring compliance with crucial safety and cybersecurity standards, such as ISO 26262 and ISO/SAE 21434. Additionally, they enhance global collaboration among distributed teams, improving operational efficiency through real-time integration capabilities that guarantee bidirectional traceability between across all work items, including requirements, risks, testing, and code. Another vital component is impact and change management; wherein automatic tracking links allow quick evaluation of the implications of any changes to requirements. These ALM tools also incorporate compliance models aligned with industry standards, simplifying documentation and verification processes related to safety and quality. Moreover, the integration of AI-powered support streamlines the development of automotive requirements, reducing manual effort and improving the accuracy of analyses and validations[48] .

3.2.2 GenAI integration in Requirement Management Tools as Future Trend

The rapid rate of technological advancement has led to significant investments in the automotive sector to remain competitive. Requirements Management has undergone substantial transformation due to technological innovation, and the adoption of Application Lifecycle Management (ALM) platforms has become essential. Nevertheless, the integration of Generative Artificial Intelligence (GenAI) into these systems remains an emerging approach and does not yet constitute a mature, market-ready solution for requirements engineering in the automotive domain.

According to *Kovair Software (2023)* [50], generative AI may represent the missing component in fully integrated, end-to-end Application Lifecycle Management (ALM) systems. It has the potential to enhance all the six stages of the ALM process. Beginning with the requirements planning phase, GenAI can reduce ambiguity in user stories and business requirements by leveraging natural language processing (NLP) to translate informal business language into precise technical specifications. During the design stage, it can assist system architects by proposing high-level design models based on predefined constraints and objectives. In the development phase, generative models can produce entire blocks of code, thereby accelerate software implementation and reduce manual effort. Moreover, in the testing and quality assurance phase, GenAI can generate synthetic test data, including edge cases that might otherwise be overlooked, thus improving test coverage and robustness. During deployment and release management, GenAI can automate the creation of deployment scripts for complex multi-environment configurations, enhancing reliability and consistency. Finally, in the maintenance and continuous improvement stage, generative AI can analyse user feedback and system usage data to identify areas for functional enhancement or performance optimization.

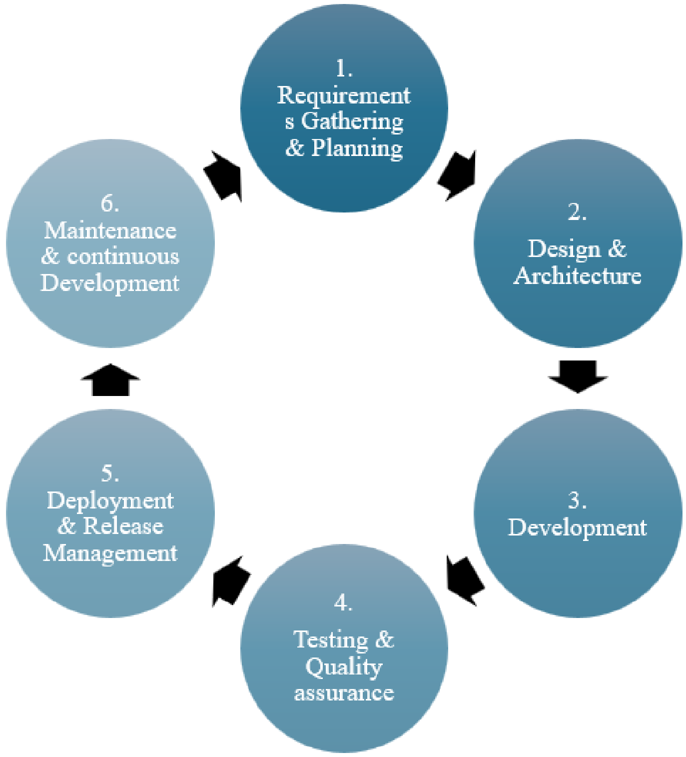

Figure 1. ALM process phases.

The integration of GenAI into ALM offers several significant benefits. First, it accelerates development by enhancing the generation of code, documentation, and test cases, ultimately reducing cycle times by 30–50%. A shorter cycle time leads to a cost reduction given that the automation of manual tasks minimizes the need for human resources and reduces the risk of errors fostering a more efficient resource utilization [49]. Generative AI fosters improved collaboration across business and technical teams, as conversational interfaces enable smoother communication and easier access to relevant information for all stakeholders. Furthermore, the quality of the software improves, as AI-driven testing and analysis help identify and fix bugs earlier in the development process, resulting in a better overall user experience. Finally, generative AI enhances scalability, enabling development to expand without a proportional increase in employees. This allows organizations to scale up their projects more efficiently, meeting growing demands without compromising on performance or quality [50].

Exploiting these tools leads also to several risks that complicate their development and market deployment. The primary risks include [42]:

- *Hallucinations:* generation of content that, while plausible, is incorrect or deviates from the original source. To mitigate this risk, companies must invest in large, high-quality datasets and powerful hardware for model training. However, a significant bottleneck remains the lack of publicly available specialized automotive datasets for fine-tuning, making the adoption of generative AI more time-consuming and challenging than initially anticipated.

- *Data Privacy*: requirements are considered highly valuable assets, and their dissemination is frequently restricted by legal mechanisms, such as non-disclosure agreements. Consequently, it is typically advisable to confine the exposure of such data within organizational boundaries. This limitation complicates the use of third-party services and external cloud infrastructures, often resulting in a preference for smaller, locally deployable models tailored to specific tasks. However, these models remain largely in the research phase, with the only viable alternative being commercial solutions, which are not adequately suited for the proper management of such sensitive data.

- *Regulatory Compliance:* in the automotive industry, compliance with industry standards, such as ISO 26262 and ISO/SAE 21434, is applied to all GenAI artifact as crucial for safety and cybersecurity. The lack of standardized automotive datasets for AI models complicates the ability to ensure compliance across different AI-driven solutions.

The potential of GenAI to streamline the entire product lifecycle is clear, yet its adoption is hindered by issues such as the lack of specialized automotive datasets, data privacy concerns, and regulatory compliance requirements. As the technology matures, addressing these challenges will be crucial for realizing the full benefits of generative

AI in the automotive industry. Robust governance frameworks and careful management of data privacy and compliance will be key to overcoming these obstacles and driving widespread adoption.

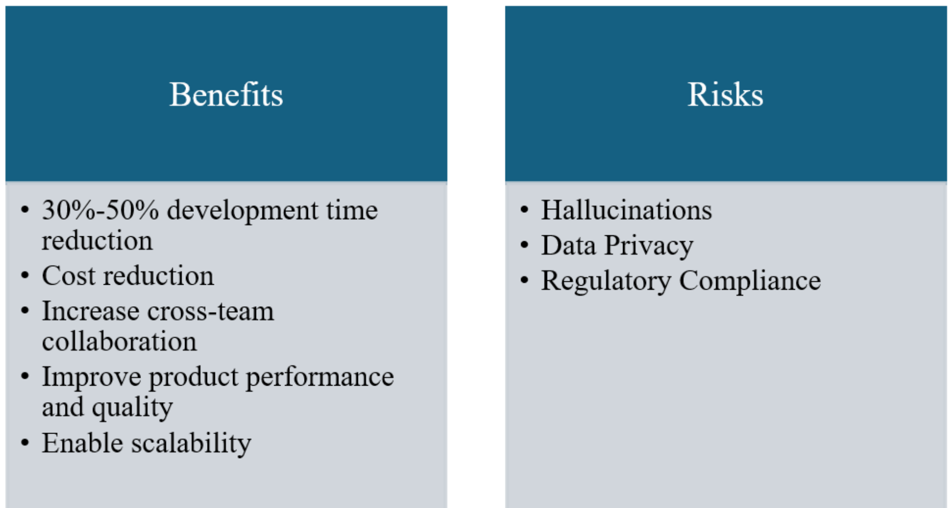

Figure 2. Benefits-Risk of GenAI-ALM integration.

4 Definition of the Product Quality Model

4.1 Mapping of requirements according to ISO/IEC 25010

4.1.1 Definition of ISO/IEC 25010

The International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC) composed the specialized system for worldwide standardization. The ISO 25010 is a part of the SQuaRE (System and software Quality Requirements and Evaluation) family of International Standards. The SQuaRE is organized in five Divisions, each of them groups together multiple similar standards. The ISO 25010 is a component of the Quality Model Division, which depicts comprehensive quality models for computer systems and software products, data, IT services and quality in-use [51].

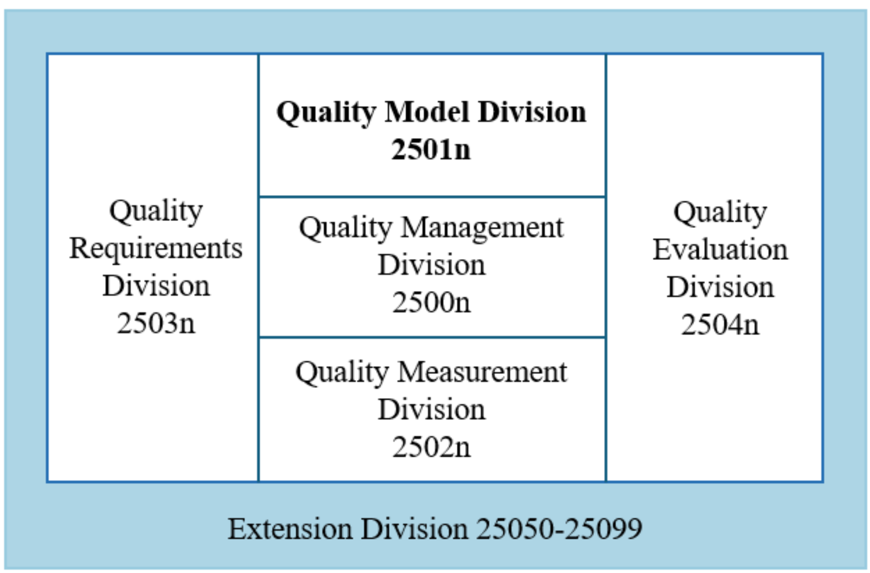

Figure 3. Organization of SQuaRE family of International Standards.

The ISO 25010 defines nine main characteristics, namely: functional suitability, performance efficiency, compatibility, usability, reliability, security, maintainability, flexibility and safety. Each of them is further divided into sub-characteristics, that provide a product quality model able to be specified, quantified and assessed. The latter is applicable to ICT (information and communication technology) and software products. The model supports product's stakeholders across various stages of the product lifecycle. It aids in defining and validating requirements, guiding system and product design toward quality objectives, and establishing corresponding testing and quality assurance criteria. Furthermore, it assists in determining acceptance conditions and measurable quality characteristics to ensure that products and information systems meet the desired standards during their development and evaluation [51].

The following table depicts all the main characteristics and their definitions.

Table 1 - The nine main characteristics of the Product Quality Model [51].

Main Characteristics	Definition
Functional suitability	Capability of a product to provide functions that meet stated and implied needs of intended users when it is used under specified conditions.
Performance efficiency	Capability of a product to perform its functions within specified time and throughput parameters and be efficient in the use of resources under specified conditions.
Compatibility	Capability of a product to exchange information with other products, and/or to perform its required functions while sharing the same common environment and resources.
Usability	Capability of a product to be interacted with by specified users to exchange information between a user and a system via the user interface to complete the intended task.
Reliability	Capability of a product to perform specified functions under specified conditions for a specified period of time without interruptions and failures.
Security	Capability of a product to protect information and data so that persons or other products have the degree of data access appropriate to their types and levels of authorization, and to defend against attack patterns by malicious actors.
Maintainability	Capability of a product to be modified by the intended maintainers with effectiveness and efficiency.
Flexibility	Capability of a product to be adapted to changes in its requirements, contexts of use, or system environment.
Safety	Capability of a product under defined conditions to avoid a state in which human life, health, property, or the environment is endangered.

4.1.2 Quality standards application to GenAI-enhanced ALM platform

To define the quality-driven conceptual framework for a GenAI-enhanced ALM tool, this project focused on four key characteristics defined in ISO/IEC 25010: usability, reliability, security and maintainability. These dimensions represent the essential pillars to determine the trustworthiness and robustness of the tool. Furthermore, they mirror the risk outlined in the previous chapter (3.2.2), which clarify why such solutions remain at the research stage and have not reached the market maturity yet.

Usability refers to the ability of the tool to translate natural language requirements into unambiguous technical specifications. This characteristic facilitates traceability across requirements, test cases and implementations, while also supporting effective collaboration among stakeholders.

Reliability concerns the operational continuity of the tool and its capacity to manage failures. It is defined by the consistent of performance, including the reduction of hallucinations and data loss. Given that, this tool will be used by multiple teams on complex project, reliability is a prerequisite for ensuring trust and adoption.

Automotive requirements frequently include sensitive or strategic information. Ensuring compliance with the security standard is therefore essential for any GenAI-integrated tool. Security encompasses access monitoring, protection against attacks or misuse and adherence to legal and regulatory frameworks.

Maintainability denotes the capability of the tool to remain aligned with evolving business models and practices. It guarantees scalability through new versions and updates, while keeping upgrade costs sustainable. Without adequate maintainability, the tool would quickly lose both its effectiveness and efficiency.

4.2 Measurable quality attributes and related indicators (KQI/KPI)

4.2.1 Definition of ISO/IEC 25023

The ISO/IEC 25023 is an International Standard part of the SQuaRE family, in the Quality Measurement Division, which defines a system product quality measurement model. The ISO/IEC 25023 provides quantitative measures for assessing system and software product quality based on the characteristics and sub-characteristics defined in ISO 25010. It integrates quality measures with guidance on their application to software and systems, omitting however specific value ranges, as they vary according to the nature of the product. The proprieties of a product are categorised on the stages of its lifecycle, presenting a distinction between internal measures, used during the development phase, external measures that can be applied at the testing and operational stages and quality in use properties, which verify the quality of the product when it is in real or simulated use. The internal measures can verify the quality issues of a non-executable tool, while the external ones evaluate the behaviour of a software when is executed. Each of the sub-characteristics described by ISO/IEC 25010, corresponds respectively to one or more measurement functions that allow their quantitative evaluation [52] .

4.2.2 Defining the main KQIs of a Product Quality Model

In the context of software quality assessment, to evaluate a product performance and enhance its quality objectives, Key Quality Indicators (KQIs) represents the perfect benchmarks, as their definition is crucial to monitor the quality of a product during the design, operation and development phases. The ISO/IEC 25023 standard provides a comprehensive framework of measurement functions that define how each quality characteristics and sub-characteristics, described in ISO/IEC 25010, can be quantitatively assessed through measurable indicators.

To establish a solid theoretical framework for evaluating a GenAI-enhanced ALM platform, this thesis derived KQIs from the primary functions associated with the main sub-characteristics relative to the four key attributes outlined in the paragraph 4.1.2, namely: usability, reliability, security and maintainability.

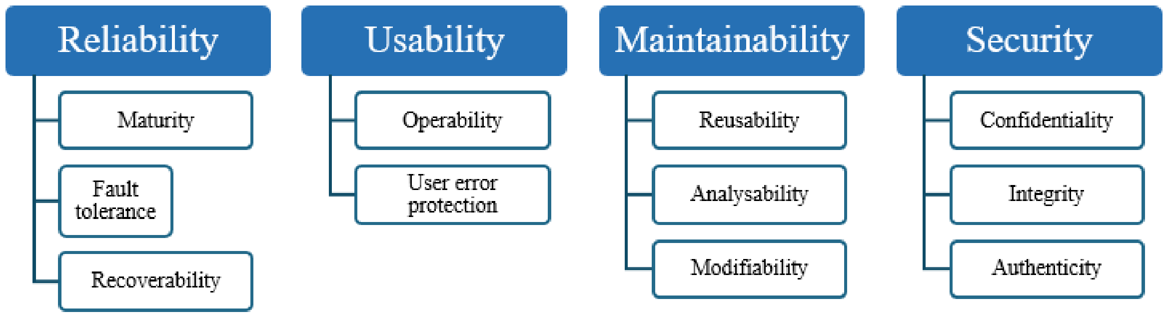

Figure 4. Selected characteristics and sub-characteristics for GenAI-ALM integration

Usability measures evaluate the system's ability to achieve specific objectives with the intended level of efficiency and effectiveness, while meeting the needs of its target users. Its *Operability* sub-characteristic is perfectly embodied by mathematical function of *operational consistency.* This KQI assesses the extent to which the system performs tasks consistently, making easy for the user maintaining effective control during operation, through ratio of interactive tasks that maintain consistent behaviour and appearance, on the number of tasks that need to be consistent. *User error protection* sub-feature evaluates the system's ability to prevent and mitigate user-induced errors. It comprises the *avoidance of user operation error* and *user error recoverability KQIs*. The first KQI quantifies the proportion of inputs protected from causing malfunction relative to all inputs that could be protected, while the second one calculates the ratio of user errors that the system is designed to recover on the total number of potential user errors during operation[52].

Reliability refers to the ability with which a product performs specified functions under defined conditions and period of time. *Maturity* sub-characteristic focuses on the capability of the system to meet reliability requirements under normal operation. This is typically evaluated through the *Mean Time Between Failure* (MTBF) and the *failure rate,* which respectively indicate the average operating time between two consecutive failures and the average number of failures observed within a specified observation period. The *failure avoidance* measurement function, part of the *failure tolerance* sub-characteristic, is used during the testing phase to determine the proportion of avoided critical failure occurrences relative to the total amount of executed test cases of fault pattern. *Recoverability* sub-feature assesses the system's ability to restore data and functionality following an interruption. Its *Mean recovery time* KQI quantifies the average time required for the system to recover from a failure, dividing the total recovery time by the number of failures. Hence, it represents the time needed to restore the software operation after each incident [52].

Security plays a crucial role in assessing software quality, as data protection is a fundamental requirement in any system. Its *Access controllability* KQI, part the *confidentiality* sub-feature, evaluates the proportion of confidential data that are protected with authorization access, subtracting the proportion of the confidential one that can be acceded without authorization, to the total. Considering the *Integrity* sub-feature, the *Internal data corruption prevention* is a relevant KQI, as it quantifies the ratio of data corruption prevention methods implemented to the total number of recommended protection methods. The strength of authentication is essential to the security of a software, and it can be assessed through the *Authentication mechanism sufficiency* KQI, part of the *authentication* sub-characteristic, which computes the ratio between the number of implemented authentication mechanisms on the total specified [52].

The *Maintainability* standard assesses the effectiveness and efficiency with which a software system remains aligned with the business model during modifications. In particular, the mathematical functions relative to the *reusability* sub-characteristic focus on the capability of integrating existing assets into new or updated components. In particular, the *reusability of assets* measures the proportion of the resources explicitly designed to be reusable on the total amount, considering requirements documents, source code modules and testing modules as assets. In the *Analysability* sub-feature, the *diagnosis function sufficiency* KQI quantifies the ratio of the implemented diagnostic functions to the total required, reflecting the aim of its class to assess the impact on a product of the changes made to its components. Finally, from the *Modifiability* sub-class, the *Modification correctness* measure*,* evaluates the proportion of correctly implemented modifications, computing the extent to which the system can be effectively updated without degrading existing product quality [52].

The definition and measurement of Key Quality Indicators (KQIs) during the product lifecycle is essential for ensuring a comprehensive approach to the overall quality assessment of the GenAI-assisted ALM tool. By continuously evaluating these indicators, it becomes possible to track the system's compliance with established quality standards in real time. Furthermore, the identified KQIs will serve as reference metrics in the subsequent phases of the project, namely: prototyping, validation and evaluation, to verify the adherence to the defined quality objectives. In the end, KQIs provide a robust foundation for informed decision-making, allowing stakeholders to identify areas for optimisation and ensuring the tool’s sustained effectiveness and relevance overtime.

Table 2 - KQIs for the evaluation of GenAI-aided ALM tool *[52]*.

Main Characteristic	Sub-characteristics	Key Quality Indicator (KQIs)	Measurement Method
Usability	Operability	Operational consistency	X = 1 – A/B A = N. of specific interactive tasks that are performed inconsistently B = N. of specific interactive tasks that need to be consistent
	User error protection	Avoidance of user operation error	X = A/B A = N. of inputs protected from causing system malfunction B = N. of inputs that could be protected from causing any system malfunction
		User error recoverability	X = A/B A = N. of user errors designed to be recovered B = N. of user errors which can occur during operation
Reliability	Maturity	Mean Time Between Failure (MTBF)	X = A/B A = Operation time B = N. of system failures occurred
		Failure rate	X = A/B A = N. of failures detected during observation time B = Observation time
	Failure tolerance	Failure avoidance	X = A/B A = N. of avoided critical failure occurrences B = N. of executed test cases of fault pattern during testing
	Recoverability	Mean recovery time	$x = \sum_{i=1}^{n} A_i / n$ Ai = Total time to recover the downed system and re-initiate operation for each failure I n = N. of failures

Security	Confidentiality	Access controllability	X = 1 − A/B A = N. of confidential data items accessed without authorization B = N. of data items that require access control
	Integrity	Internal data corruption prevention	X = A/B A = N. of data corruption prevention methods implemented B = N. of data corruption prevention methods recommended
	Authenticity	Authentication mechanism sufficiency	X = A/B A = N. of authentication mechanisms provided B = N. of authentication mechanisms specified
Maintainability	Reusability	Reusability of assets	X = A/B A = N. of assets which are designed and implemented to be reusable B = N. of assets in a system
	Analysability	Diagnosis function sufficiency	X = A/B A = N. of diagnostic functions implemented B = N. of diagnostic functions required
	Modifiability	Modification correctness	X = 1 – (A/B) A = N. of modifications that caused an incident or failure within a defined period after being implemented B = N. of modifications implemented

5. GenAI Integration and Roadmap for ALM Tool Development

5.1 State of the Art and Existing Gaps

As previously discussed in Chapter 3, the automotive sector is at the forefront of technological investment from many points of view, from software development to production, up to infotainment. Another crucial field is the Requirements Management, which is not only fundamental but also very complicated, given the high number that each automotive component must satisfy. For this reason, in recent years, technological advancements have tried to address this issue, aiming for a solution that meet both customers' and workers' needs, respect the passengers' safety measure, the data protection, and simultaneously optimize time management while providing added value.

5.1.1 Research gaps in GenAI-Driven Requirements Management

As highlighted by H. Cheng *et al.* (2025) [53], there is a research gap regarding the management of requirements using LLMs in the automotive sector. To fill this gap, the H. Cheng *et al.* (2025) [53] study offers a comprehensive analysis, from 2019 to 2024, on the current stage, challenges and future trends of GenAI in Requirements Engineering (RE). They formulate four research questions regarding respectively, the current research trends, the approaches, the quality of these studies and the challenges in the application of GenAI in RE. The first one outlined that most of the publication (92.4% of 105 papers analysed) were full papers, reflecting a trend toward comprehensive research. However, Requirement Management reached only 5.4 %, confirming the limited attention paid to this field. Although it appears that the industrial application of GenAI in Requirements Management is growing, research in this area remains limited, highlighting the lack of attention to this rapidly developing field. The second question, regarding the model selection, points out an adoption rate equals to 67.3% for GPT series models, especially of GPT-3.5 and GPT-4, with only a 7.4% in open-source alternatives such as LLMA, mirroring a limitation in exploring diverse approaches [53]. The outcome of the N. Petrovic *et al. (2025)* [42] study corroborates these results, showing the dominant position of the commercial GenAI solutions in the market, despite their inability of managing the automotive requirements. The third research question, based on quality assessment, shows high academic standards, with clarity of objectives and methodological rigor; however, improvements are needed in data analysis and discussion of limitations for more reliable and contextualized applications. The H. Cheng *et al.* (2025) [53] last question focuses on the challenges of applying GenAI in RE; its outcomes highlight a strongly correlated triad composed by interpretability, reproducibility and controllability. These are added to the previous risks discussed in Chapter 3.2.2, namely: hallucinations, data privacy and regulatory compliance. Concerns related to ethics, security and bias emerge as governance challenges that are still barely explored, while computational costs and response times represent practical obstacles to large-scale adoption. Hence, they advise the adoption of integrated approaches with validation, bias control and anti-hallucination protocols, ensuring human supervision, especially in critical and real-time settings. Additionally, the N. Petrovic *et al. (2025)* [42] research named other issues of the integration of GenAI in RE, such as finding powerful hardware capable of supporting high computational demands and the lack of publicly available specialized automotive datasets for fine-tuning, due to data protection policies and restrictions on sharing sensitive information [42].

The complexity of requirements, the shortage of professionals and data leakage outside the company, are addressed by the study of Y. Uygun and V. Momodu (2024) [54] which proposes Natural Language Processing (NLP) Question-Answering (QA) systems. It utilizes advanced language processing techniques to recognize ambiguous phrases, eliminate linguistic uncertainty and respond to natural language queries, thus enhancing Requirements Management. Furthermore, they overcome the issue of data dispersion and the lack of LLMs for small/medium-sized businesses and independent researchers through locally executable versions of open-source ChatGPT substitutes, which are smaller but offer equivalent or even better performance than ChatGPT and GPT-4. The QA retrieval solution demonstrates that massive language models can significantly reduce the time and effort required for requirements analysis in automotive engineering, especially in scenarios with complex and lengthy documents.

These models are particularly useful for automating requirement extraction and enhancing collaboration among engineers, vendors, and customers by efficiently answering queries related to specifications and constraints [54].

5.1.2 GenAI-ALM: an industry-driven initiative

As demonstrated by the first research question of H. Cheng *et al.* (2025) study [53], Requirements Management is one of the least explored areas in research. This gap leads to deficiencies in the application of GenAI in industry, particularly regarding the preliminary evaluation of the quality and performance of these tools prior to their adoption within the corporate environment. In fact, the integration of a GenAI tool into an ALM platform is a purely industry-driven initiative, for which there are no targeted academic studies yet. This mirrors how these tools emerge from the need to optimize complex processes, such as the Requirements Management in the automotive sector. Moreover, the absence of a mature and stable market is indicative of the innovative nature of the development and application of these technologies. Indeed, can only be found pioneering examples of collaborations among companies, such as Volkswagen, Microsoft and PTC (Parametric Technology Corporation), as well as HCLTech, Siemens and AWS (Amazon Web Service).

The first one, integrates Microsoft Copilot, into PTC's Codebeamer, a widely adopted ALM platform enhancing the productivity of Volkswagen's software engineers by automating the drafting, revision, validation of requirements and test cases with an estimated 20-40% of time saving [55]. The HCLTech ALMate solution for Siemens *Polarion*, powered by AWS AI infrastructures embeds GenAI at multiple levels of the ALM lifecycle. It supports the automated generation of both requirements and test cases, ensuring traceability across the entire lifecycle and includes intelligent summarization and translation features, which improve global team collaboration and reduce delays caused by miscommunication [56]. Both solutions aim to optimize requirements management through GenAI. The Volkswagen–Microsoft–PTC case primarily focuses on improving the user experience by embedding AI assistance within established workflows. In contrast, the HCLTech–Siemens–AWS collaboration offers a more modular, infrastructure-driven architecture designed for deeper AI integration, automation, and long-term scalability across the product lifecycle. This approach represents a significant step forward into the evolution of GenAI applications within the ALM domain.

The gaps identified in the academic literature together with the immaturity of the current market for GenAI integration into ALM platforms, underline a tangible opportunity for innovation. In response to this need, this thesis proposes a hypothetical GenAI-powered ALM tool designed to address the complexities of Requirements Management in the automotive industry. The tool aims to fill the current technological void and respond to the emerging demands of an industry where GenAI applications remain in an early and largely exploratory phase of development.

5.2 Make-or-Buy evaluation

The Make-or-Buy analysis represents a critical decision-making process used to determine whether to develop a solution internally or acquire it from external providers. This evaluation is particularly relevant when assessing the integration of Generative AI into Application Lifecycle Management (ALM) platforms, especially within the highly demanding and regulated automotive sector. It enables a company to compare the strategic, technical, and economic implications of adopting an existing GenAI-enhanced tool versus developing a customized solution. This analysis is usually based on the cost and profitability computation of both scenario and the investigation of the results.
As presented by the M. Arora and A. Kumar study [57], there are drivers and concerns to both make and buy scenarios. On one hand, it identifies as factors supporting insourcing: quality, cost, design, lead times, warehousing expenses, transportation, political, environmental and social reasons. On the other hand, all the expenses relative to labour, operation, managing, purchasing, capital and material are the relative concerns. Regarding the drivers of outsourcing, there are: suppliers, lack of suppliers' expertise and all the

aspects that the providers sell at lower cost. Instead, the factors in contrast with buying are all the transportation, go-down and management expenses [57].

M. Arora and A. Kumar study [57], divided drivers and concerns of outsourcing into five different categories, such as: Core business, Reliance on third parties, Cost dependency, Size of requirement and workload, Lack of expertise and new product. If an activity is not central to the company's value proposition, outsourcing may be strategically valid. However, this choice can lead to a loss of core activities and a weakening of critical capabilities. Hence, organizations must carefully assess whether outsourcing may erode strategic assets over time. The dependency on third suppliers raises concerns regarding quality, reliability, and operational continuity. It requires a robust supplier selection process, with attention to technical competence, reliability, and long-term trustworthiness. Unreliable or unqualified suppliers pose serious risks and should be excluded from consideration. Lastly, in sectors handling sensitive data or high compliance requirements, supplier trust becomes a central element of the decision-making process. From a financial perspective, outsourcing can convert fixed costs into variable costs, offering increased flexibility in resource allocation. However, this may come at the expense of internal control. Furthermore, there are often hidden costs associated with procurement contracts, such as raw materials, service levels, or integration overheads, that must be properly accounted for. Without a comprehensive and well-structured cost-benefit analysis, the organization risks underestimating the true economic impact of the outsourcing decision. The scale and variability of requirements also influence the make-or-buy strategy. For small-volume or highly variable workloads, outsourcing may be more efficient. Internal development, by contrast, involves greater human resource allocation and may be less flexible in adapting to demand fluctuations. Additionally, sales volatility and changing market conditions can significantly affect the cost-effectiveness of either option. One of the most critical drivers is the availability of internal expertise and new product. When organizations lack the necessary skills, production facilities, or operational capacity, outsourcing may appear to be a viable solution. However, choosing to buy also implies the need to disclose proprietary technology, which can compromise design secrecy and intellectual property protection. For this reason, when developing innovative products with key proprietary components or strategic parameters, organizations often prefer to retain control through internal development [57].

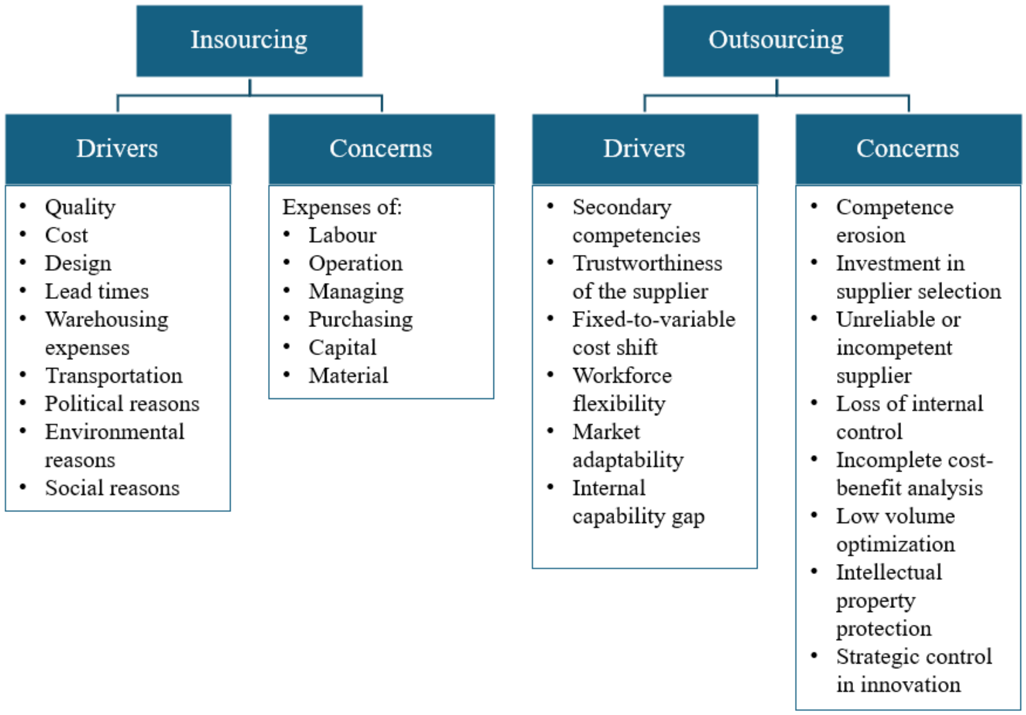

Figure 5. Make or Buy Drivers and Concerns.

Considering the automotive sector, its core business concentrated around the production and commercialization of vehicles. Hence, for companies operating in this field, investing significant resources to internally develop an ALM platform integrated with a GenAI-based requirements management tool could prove burdensome. Such project may be considered a secondary competency, diverging from the core business of the automotive industry. It would require access to powerful hardware capable of handling high computational demand, specialized personnel, and the availability of fine-tuning datasets compliant with stringent data privacy regulations, as previously discussed.

On the other hand, purchasing a commercial solution entails additional consideration. While it may facilitate a shift from fixed to variable cost, it also introduces risks such as the loss of internal control, incomplete cost-benefit analysis and dependency on the trustworthiness of the supplier. The latter becomes particularly important when delegating not only operational tasks but also sensitive data, which raises concerns about intellectual property protection and strategic control in innovation. Although no reliable data are currently available to support a quantitative make-or-buy decision, a logical cost-benefit reasoning suggests that internal development would require significantly greater investment. This is due to the high demands in terms of infrastructure, datasets and specialized personnel, as previously discussed. Indeed, the collaboration among Volkswagen Microsoft e PTC is a perfect example of outsourcing, in which Volkswagen leverages the external competencies of Microsoft and PTC to manage requirements within its development processes.

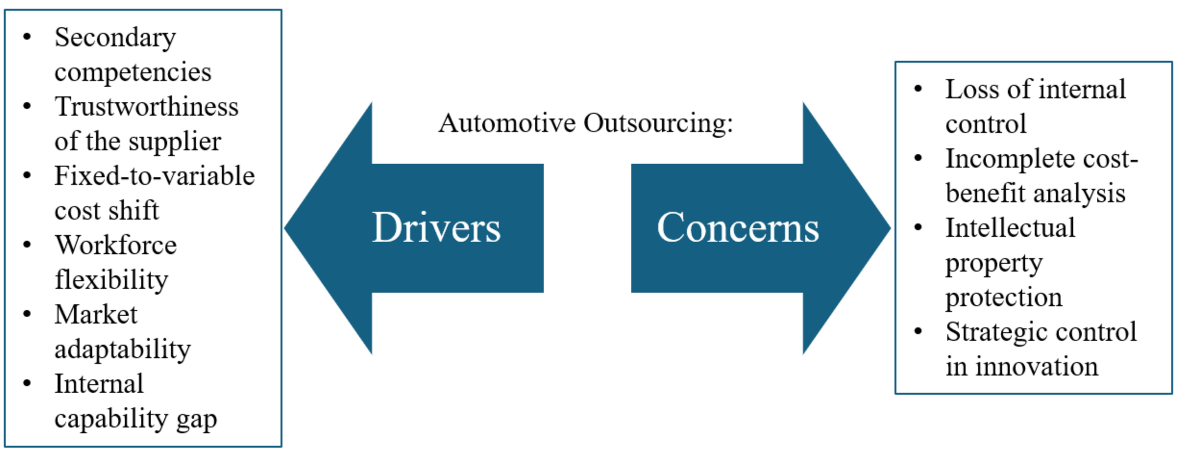

Figure 6. Automotive Outsourcing: Drivers and Concerns.

5.3 Requirements Management tool assessment.

ALM (Application Lifecycle Management) platforms are fundamental for managing requirements in the automotive sector, especially to ensure compliance with rigorous standards such as ISO 26262 and Automotive SPICE. Requirements Management software provides, through a centralized archive, a single source of information for all interested parties and allows the documentation and traceability of all changes, ensuring compliance with quality standards. This tool promotes collaboration between cross-functional teams by providing features such as versioning, comment threads, and approval workflows. These features allow all members to stay aligned and contribute effectively to project requirements. Moreover, it minimizes errors and manual work through automation, improving project efficiency and time management [58].

Hence, according to the Guidelines proposed by *Visure* [58], the main factors and characteristics to evaluate to select the right Requirements Management Tool are:

1. Requirements Traceability: to ensure complete end-to-end requirements traceability, these tools must enable teams to link requirements across the entire project lifecycle, providing visibility into dependencies, facilitates compliance, and simplifies change management.

2. Traceability matrices: to link requirements to design, testing and implementation, these tools should provide real-time updates to track changes, their impact and ensure compliance with regulatory and quality standards.

3. Version control for requirements: these tools should allow easy monitoring, comparison and recovery of previous versions for branching and merging options for parallel development.

4. Collaboration and review tools: real-time collaboration and shared editing, requirement comments and annotations, automatic notifications, and approval workflows.

5. Integration with other project and Requirements Management software: essential for improved project visibility. A company should look for a two-way integration with testing and project management tools, data synchronization to avoid duplication and errors import/export functionality for smooth transitions between tools.

5.3.1 Strengths and weaknesses of the most relevant tools for the automotive industry

Considering the study carried out by *Visure* [59] , which underlining the main advantages and disadvantages of the ten best Requirements Management software of the 2025, this thesis mainly focuses on the most relevant tools for the automotive sector, such as *IBM DOORS Next, Visure Solutions ALM Platform* and *Codebeamer ALM*.

IBM DOORS Next is the successor to IBM DOORS, it was designed by IBM to define, track, analyse, and manage complex product and system requirements throughout the entire development lifecycle. It provides a structured and organized environment for capturing, documenting and analysing requirements, using modules or collections. The platform incorporates version control functionalities, which facilitate the monitoring requirements evolution over time while mitigating potential conflicts. The tool's customization is a fundamental characteristic, enabling alignment with specific organizational needs and workflows, thereby ensuring versatility and flexibility. It guarantees the integration of requirements management into the broader development process, thanks to both internal linkages between requirements at different levels and external connections with other tools. The system offers robust reporting and analytics capabilities, allowing stakeholders to generate customized reports and dashboards. Regarding the disadvantages characterizing IBM DOORS Next, these include the complexity of its functions, which may require extensive training and insertion; the high cost of licensing and maintenance that makes it less suitable for small companies or projects; and consequently, the significant number of resources in terms of both hardware requirements and specialized personnel [59].

Visure Solutions ALM Platform, focuses primarily on Requirements Management, followed by risk, test and artefacts administration capabilities, in order to provide a unified environment that ensures standards compliance during the product lifecycle. The platform delivers a comprehensive system that enables organizations to analyse, document and manage requirements, addressing all aspects of Requirements Engineering. It has gained considerable recognition in highly regulated sectors, particularly the automotive industry, attributed to its robust support for regulatory compliance requirements. The platform's end-to-end traceability ensures that modifications in one domain are systematically reflected and tacked across the entire project, enhancing transparency and accountability. Visure's integration with other ALM tools enables teams to leverage their existing technological infrastructure while augmenting requirements engineering capabilities. Its customization and flexibility features ensure adaptability to specific project requirements and organizational contexts. This tool, through notifications, discussion threads and commenting, mitigates miscommunication and contribute to higher-quality outcome. Additionally, it supports the generation of customized reports and dashboards, providing stakeholders with tailored analytical insights into project status and requirements quality metrics. However, Visure Solutions ALM could represent an overly designed solution for organizations engaged exclusively in short-term projects with minimal project dependencies. Additionally, the platform's broad feature set and compliance-oriented architecture may prove unnecessarily complex and require many resources for projects that do not require rigorous traceability, regulatory compliance, or comprehensive lifecycle management capabilities, with the risk of obtaining suboptimal cost-benefit ratios [59].

CodeBeamer embeddes an extension of ALM functionalities with product line configuration capabilities and provides unique configurability for complex processes.

The platform delivers wide ALM support encompassing the entire development continuum, from requirements management through development, testing, and release management, thus ensuring continuous integration of requirements within the broader development ecosystem. Its robust requirements management enables teams to efficiently capture, document, track, and prioritize requirements, ensuring clear definition and facilitating traceability throughout the project lifecycle. A distinguishing strength of CodeBeamer lies in its sophisticated traceability and impact analysis capabilities, which allow users to establish and maintain bidirectional traceability links between requirements, design elements and test cases, ensuring regulatory compliance. The platform's high degree of customization and adaptability accommodates diverse project types and methodologies, enabling organizations to tailor the tool to their specific requirements and workflows. Furthermore, CodeBeamer's seamless integration with widely adopted development and testing tools, ensures cohesive requirements management within the development process. Advanced reporting and analytics functionalities provide project managers and stakeholders with customized dashboards and data-driven insights that support informed decision-making and effective project monitoring. However, several limitations warrant consideration. The platform's many features can make it difficult to learn, requiring significant training and onboarding investments to fully leverage its capabilities. CodeBeamer can be resource-intensive regarding hardware requirements and system performance. Licensing and maintenance costs are relatively significant, particularly for teams or organizations with large requirements management needs, rendering the solution less economically viable for smaller organizations or budget-constrained projects [59].

IBM DOORS Next, Visure Solutions ALM, and *Codebeamer* share a common foundation in supporting essential requirements management, traceability, compliance, and lifecycle integration tasks. These platforms demonstrate strong alignment with the operational and regulatory demands characteristics of the automotive industry, delivering robust functionality, high configurability and advanced integration capabilities. Nevertheless, they also have similar limitations across several dimensions: high licensing and maintenance costs, requirements for specialized training and onboarding programs, and significant demands on both hardware infrastructure and specialized personnel resources. These constraints can impede their adoption in smaller or less mature organizations. Although the limitations of *Visure Solutions ALM* are presented implicitly within the available literature, given the source characteristics and industry positioning, the analytical evidence suggests that its disadvantages substantively parallel those observed in the other ALM platforms examined.

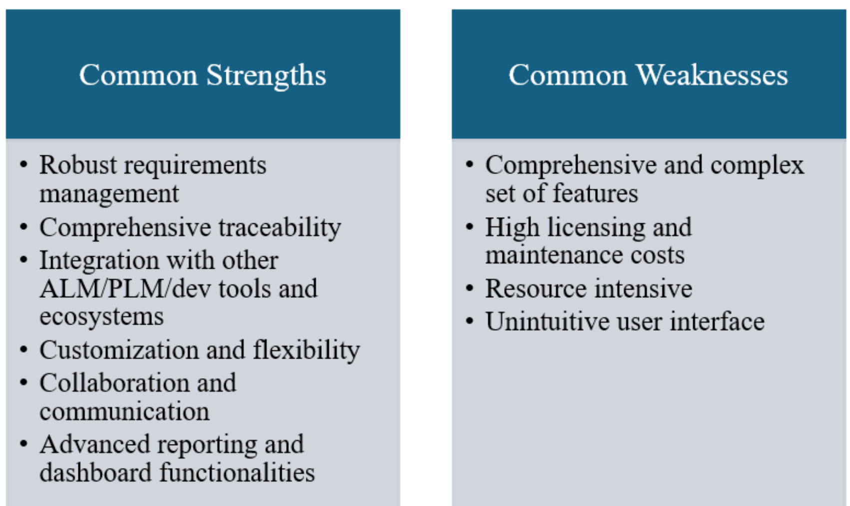

Figure 7. Common Strengths and Weaknesses of the ALM analysed.

5.3.2 Platform suitability for GenAI adoption

The integration of Generative Artificial Intelligence into enterprise platforms, such as ALM and requirements management tools, represents a promising and complex innovation opportunity. The platforms strengths already discussed constitute essential prerequisites for such integration. The study conducted by Haki et al. (2024) [60], based on empirical evidence from the Salesforce ecosystem, considered a prime example of an enterprise platform and a pioneer in GenAI adoption, demonstrates that effective GenAI integration into any business platforms, requires substantial modifications at three platform levels: capabilities, architecture, and governance.

At the architecture level, GenAI acts as a “cross-layer component” that enhances natural language processing, supports domain experts, and automates repetitive, low-value tasks. The integration introduces new architecture components such as prompt engineering pipelines for GenAI model interaction, trust layers, which simulate protected versions of real data to mitigate privacy risks, and zero-retention policies ensuring that prompts and input data are not stored by GenAI models.

Considering the governance stage, the introduction of GenAI reshapes organizational responsibilities, licensing models, and data ownership schemes, requiring clear frameworks for ethical use and explainability. To support integration, the platform providers are also expected to deliver educational and consulting packages aimed at maximizing business value, while fostering partnerships with both emerging and established GenAI providers to expand their technological and market reach.

Finally, the platform capabilities side, wherein all the existing features offered are extended to the GenAI tool, and generic and industry-specific use cases must be offered together with “specialised GenAI studios” including modules for model configuration, application development, and prompt engineering, which are essential to ensure adaptability and user control in enterprise environments [60].

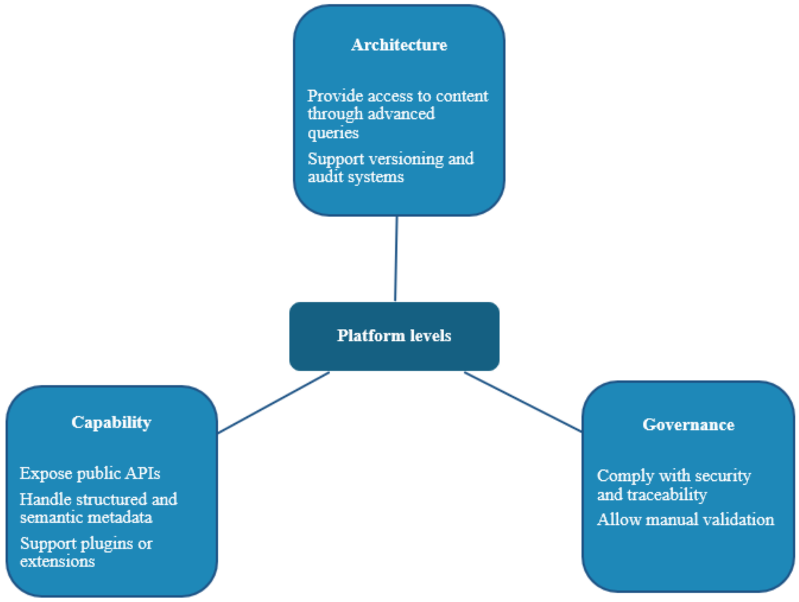

Figure 8. Platform levels and their modifications.

5.3.3 Quality driven comparison of ALM platforms for GenAI integration

Considering the features outlined above and the ISO/IEC 25010 quality standards previously covered in Chapter 4, all three platforms, *IBM DOORS Next, Visure Solutions ALM, and Codebeamer*, show both strengths and weaknesses about the integration of a GenAI tool. Concerning *reliability*, their general robustness and traceability grant a strong foundation. However, the complexity and resource intensive nature of these tools, can lead to a potential risk of performance degradation, managing the LLMs workload. Features such as, customization and flexibility, collaboration and review tools, characterized all the three platform and are valuable enablers of *usabilty*. Though, the Codebeamer's considerable learning curve and IBM DOORS Next functions complexity, together with the unintuitive user interface of both, are direct obstacles to usabilty. As for maintainability, Codebeamer stands out for its product line configuration capabilities (PLE) and unique configurability for complex processes, which suggest maximum architectural flexibility to adapt the workflow to GenAI integration [61]. Likewise, IBM DOORS Next benefits from integration with the IBM

Jazz platform and adherence to the OSLC (Open Services for Lifecycle Collaboration) standard, improving interoperability and enabling robust data exchange across the engineering lifecycle, crucial for long-term GenAI maintainability [62]. Security represents a critical challenge for all three platforms, given the risks of GenAI discussed in the section 5.1.2. Although all platforms operate in regulated sectors (ISO 26262, ASPICE), demonstrating a solid security foundation, Visure Solution ALM distinguishes itself through strong regulatory compliance and end-to-end traceability [59]. The latter provide a solid governance framework for managing LLM-generated outputs, ensuring traceability and verifiability. Nonetheless, the limited documentation available on Visure's system architecture suggests reduced flexibility for custom GenAI integrations.

Comparing the robustness and maintainability, IBM DOORS Next is the most suitable, ensuring strong integration into a broad ecosystem. Its use of the OSLC standard offers access to rich metadata, essential for GenAI, and facilitate interoperability while minimizing long-term integration risks. In contrast Codebeamer, is ideal for businesses aiming to leverage GenAI for highly specific and complex processes, such as product variant management, thanks to its flexible and modern architecture. However, this leads also to an increase in complexity and a steeper learning curve. Lastly, Visure Solutions ALM is the preferred choice in highly regulated environments where traceability, verifiability, and governance take precedence over configurability or openness.

In conclusion, the decision to integrate GenAI into an ALM platform should be preceded by a clear definition of the intended use case and a risk-benefit analysis. Only by defining the strategic priorities and assigning weights to evaluation criteria according to these, organizations can identify the platform that offers the most aligned and sustainable path for GenAI integration.

5.4 Hypothetical GenAI-enhanced ALM Tool: Architecture and Functional Scope

The conceptualization of a hypothetical ALM tool enhanced with GenAI arises from the challenges relative to the Automotive Requirements Managements, faced by MCA Engineering S.r.l. In this highly regulated and technologically advanced domain, the need for technological primarily originates from a operational and business obligations. However, the limited academic research on this specific integration, leaves the industrial implementation of such tools without a consolidated theoretical basis, evidencing the early and underdeveloped stage of GenAI applications in ALM platforms.

5.4.1 Functionalities

The integration of GenAI within an ALM system may represent the pivotal component to improve all phases of a ALM process and hence of the production lifecycle. Beginning with the requirements phase, GenAI can reduce ambiguity in user stories and technical specifications through advanced natural language processing techniques. The requirements traceability is facilitated by the automatic translation of the informal company language into structured and machine-readable formats. It ensures full visibility and version control, supporting effective monitoring, branching and restoration of historical requirements changes. During the design phase, GenAI can support system architects by generating high-level models based on predefined constraints and design objectives. It enhances collaborative design and review tools, enabling real-time co-editing, comments, approval workflows and notifications for early system prototyping. In the testing and quality assurance phase, GenAI contributes by generating synthetic test data and borderline cases that increase the coverage and robustness of the tests. In the deployment phase, it can automate configuration scripts for complex multi-environment configurations, improving reliability and integration capabilities with external systems. This ensures interoperability and data synchronization between tools such as testing and project management software.

Finally, during maintenance and continuous improvement phase, GenAI can analyse user feedback and performance data to identify functional inefficiencies and propose targeted updates or architectural optimizations, thus fostering data-driven, adaptive evolution of the system overtime [50], [58].

5.4.2 Logical Architecture

According to the analysis conducted during this thesis, the most effective architecture detected for a GenAI-powered ALM platform, is structured around four fundamentals' modules: the ALM Core, the GenAI module, a secure database and an Application Programming Interface (API), each of which plays an essential role in enhancing the overall functionality and performance of the ALM process.

The necessity of an ALM Core is described by the work of *D. Safaei and K. Haki (2024)* [60], that explains the evolution of enterprise software platforms over the last decade. In recent years, these platforms have gone from being static products to dynamic and innovative ecosystems, where the core enterprise software, is provided by the platform owner, while the partner complementors improved the platform's technology and functionalities through complementary customised extensions. This modular approach ensures both scalability and adaptability in response to emerging enterprise needs [60].

The integration of a GenAI module, previously explained and discussed in H. Cheng *et al.* (2025) [53], constitute the technological disruption aimed at overcoming the key challenges in Requirement Management, thanks to the natural language processing, automated documentation and predictive analysis. Despite its benefits, this tool introduces challenges, such as hallucination, data privacy risks and compliance issues, which can be settled by improving the specialized databases. The latter in fact, enable GenAI tools to process sensitive information in accordance with the stringent regulatory frameworks that characterize the automotive sector. Choosing between commercial and open-source GenAI models must be aligned with the company's infrastructure and risk tolerance. On one hand, commercial LLMs such as ChatGPT 4 and Gemini are widely adopted across industries, even if their use entails data exposure risks that may be incompatible with the stringent privacy requirements of automotive applications. On the other hand, open-source alternatives offer greater control but demand a non-negligible use of resources, in infrastructure and employee training. Nonetheless, as *Uygun et al. (2024)* study [54] demonstrated, emerging “small LMs” can offer a scalable alternative that balances performance and resource intensity.

The presence of a secure database serves a dual purpose. First, it acts as a preventive mechanism to avoid, particularly in cases of commercial AI are used, unauthorized access by third-party providers [60]. Second, it enhances the GenAI reliability, minimizing the hallucination problem, guaranteeing its operability on accurate and confidential company information. This is reached through the Retrieval-Augmented Generation (RAG) technique, that enables GenAI to achieve the external knowledge bases, such as these specialized databases [63]. As shown in the *Uygun et al. (2024)* study [54], the vector database, where complex data, such as words or images are represented as vectors, enables meaningful comparisons between user prompts and internal data, enhancing output relevance and precision.

Application Programming Interfaces (APIs) in the context of GenAI integration into a ALM platform, plays a role of extreme importance. APIs define standardized set of rules, protocols and tools that allow software applications to communicate with each other, exchanging data and services [64]. They allow the GenAI module to access platform data, such as requirements, models, test cases, and return generated outputs in an interoperable way, preserving the tracking, coherence and governance of the contents. Industry leaders have already implemented robust API architectures to support this type of integration. For example, IBM employs the Open Services for Lifecycle Collaboration (OSLC) REST APIs in its DOORS Next platform, enabling standardized lifecycle data exchange [65]. Similarly, PTC’s Codebeamer exposes REST APIs to support their high degree of configurability and to facilitate the integration of custom GenAI modules [66]. The ability to access specific data securely and bidirectionally via API is a foundational prerequisite that allows generative

Artificial Intelligence to act effectively and in compliance with regulations within an ALM workflow.

In the automotive sector integration between the ALM platform and the GenAI Module follows a rigorous bidirectional flow, essential for ensuring data quality (RAG) and traceability. The process begins in the Core ALM, where the user requests a GenAI action (e.g. generating a test case). Before sending the message to the GenAI module, the system, via the APIs, queries the ALM's secure Database. This preliminary consultation is fundamental for the RAG technique, wherein, by recovering the specific proprietary context such as previously approved specifications, domain-specific terminology, or regulatory constraints, the system creates an enriched prompt, which is forwarded to the GenAI module and minimise the risk of hallucinations [54]. The generative model produces the output solely based on this curated context, ensuring that responses are grounded in verifiable enterprise knowledge [67]. This architecture prevents unauthorized data exposure, particularly in the case of commercial GenAI models.

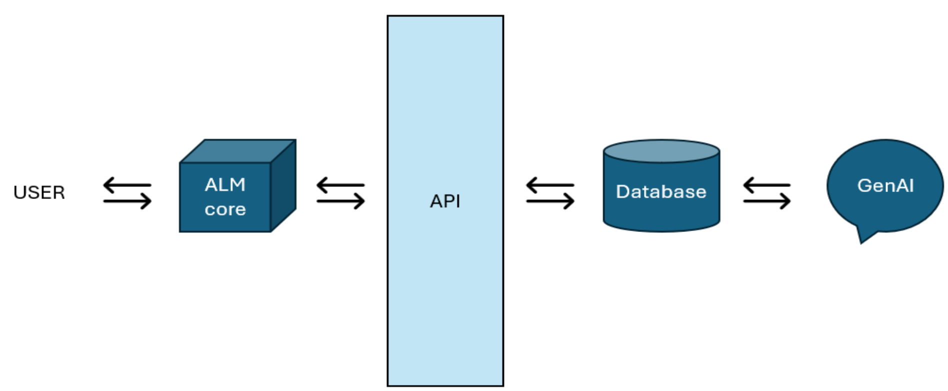

Figure 9. Logical Architecture of the GenAI-ALM platform.

5.4.3 Objectives

The development of an ALM platform integrated with Generative AI emerges as a strategic response to the increasing complexity of Requirements Management in the automotive sector. The primary objectives are:

- Development time reduction: automating time-consuming tasks, such as requirement translation, documentation and test generation, savings from 10 to 40 % of development time, in the most impactful use cases [37].

- Quality and compliance improvement: GenAI, acts as a quality controller, analysing regulatory standards and historical project data to automatically generate detailed, compliant requirement specifications. In this way, it enables a 44 % improvement in productivity, while increasing code efficiency and reliability [37], [60].

- Cross-team collaboration: GenAI supports developers and non-technical experts through low-code or no-code interfaces, thus facilitates real-time assistance, documentation, and explanation of logic across teams [37].

5.5 Project Roadmap

This section outlines a Project Roadmap to support the conceptualization and future development of a GenAI-based Application Lifecycle Management (ALM) platform. This tool is designed to address the specific complexities of the automotive industry while enabling quality evaluation based on the ISO/IEC 25010 standards, specifically: maintainability, reliability, security and usability. It is divided into four sequential phases, namely: Requirements Analysis, Prototyping, Validation and Evaluation, Scaling and Industrialization. All of which, is designed to guide the integration process, from the initial collection of requirements to a scalable implementation at the industrial level, highlighting the main results, objectives and methodological considerations. Each phase of the roadmap is associated with a Quality Gate (QG) planned to verify the achievement of quality objectives and to ensure that identified risks are solved before advancing to the subsequent stage. The quality gates (QG1-QG4) ensure traceability between requirements, validation criteria and corrective actions defined through DFMEA analysis (Paragraph 6.2).

5.5.1 Requirements Analysis

This phase focuses on defining strategic objectives, use cases, risks, and qualitative characteristics that guide the development of the GenAI-enhanced ALM tool. It aims to assess the current internal needs of MCA Engineering S.r.l. and the broader automotive context, emphasizing the challenges, opportunities, and expected impacts of integrating Generative Artificial Intelligence into Requirements Management processes.

Objectives

The development and implementation of GenAI-enhanced ALM platform aim to overcome the challenges currently faced by engineering teams at MCA, including the time-consuming and repetitive nature of requirements drafting and verification, the management of large and heterogeneous input sources, often exceeding 10,000 for each project, and the need for greater standardization and objectivity in the formulation of technical specifications.

The integration pursues the following objectives:

- Reduction of complexities within Requirements Management process.
- Innovation of internal workflows.
- Optimization of operational costs.
- Strengthening of MCA Engineering’s technological competitiveness.

Use cases

Considering the company's current practises and emerging GenAI capabilities, the most relevant and achievable use cases identified are:

- Generation of new requirements from user input or NLP.
- Automatic test case generation from structured requirements sets.
- Automatic summarization and refinement of existing requirements documents.

Risks

During the first phases, a Design Failure Modes and Effects Analysis (DFMEA) is performed. It is described in the 6.2 section of this thesis, with the objective of foreseeing and determining the main Failure Modes of each component, with the relative effects, causes and corrective actions. The latter will be implemented, in the subsequent stages, as preventive measures of the forecasted risks.

Available data and resources

Internal data sources in MCA Engineering, comprises a corporate repository of historical requirements that can be leveraged for model training and validation, feedback from active consultants and project managers, who represent the primary end users of the system. At the current stage, no formal list of GenAI-specific requirements has been defined, as it will emerge iteratively through user consultation and prototype evaluation.

Stakeholders

The stakeholders involved are:

- Business Managers, focusing on cost optimization and strategic impact.
- Delivery Managers, defining project timelines.
- Consultants and technical users, who will interact directly with the system and provide operational feedback.
- Clients: will assess the market value and potential commercialization of the solution in the future steps.

Regulation compliance

In the context of GenAI adoption, a crucial aspect concerns the storage and handling of customer data. At MCA Engineering, the implemented models performing offline inference, meaning that all computations are executed locally on the company's hardware, without transferring data to external servers via APIs. This approach ensures a higher level of data protection and regulatory compliance, as all sensitive information remains within the company secure environment.

Timeline

The four main phases foreseen by this preliminary roadmap are:

1. Requirements Analysis: current stage.
2. Prototyping: short-term implementation and internal testing.
3. Validation and Evaluation: structured quality and risk assessment.
4. Scaling and Industrialization: integration and market readiness.

Specific timeframes will be refined once the prototype architecture and resource are finalized.

Key Quality Indicators

The KQIs, described in 4.2.1 paragraph, were identified considering the main mathematical functions described in ISO/IEC 25023, corresponding to the four key features (maintainability, security and usability) and their most appropriate sub-features defined by the ISO/IEC 25010 standard.

Project's purpose

The project is initially outlined as a research and development initiative, aimed at internal use to validate GenAI integration within ALM tools. In the medium term, the objective is to scale and commercialize the solution as a service offering for external clients. No dedicated funding has been allocated yet, as a complete project proposal must be formalized before applying for internal or external financing programs.

QG1 – Requirements Approval Gate

To proceed to the subsequent phase, the QG1 must verify the completeness of the requirements, the KQIs alignment with both ISO/IEC 25010 and ISO/IEC 25023 and the stakeholders' approval. After the analysis of all the components defined above, the experts will decide whether to make this project concrete and so, approve funds and resources.

5.5.2 Prototyping

The prototyping phase, will focus on the development of a proof-of-concept (PoC) based on the logical architecture defined in this thesis (Paragraph 5.4), aimed at demonstrating the feasibility and potential benefits of the GenAI-powered ALM platform. This stage begins with the selection of technological solutions, based on the logical architecture already described, which comprises: ALM core, GenAI module, Secure Database and API layer. It identifies which language model use (open-source or commercial), data management strategies and integration APIs. Moreover, this phase focuses on the possible risks linked to the data security, implementing the corrective action defined during the DFMEA in the previous step, namely: zero-retention policy and data governance to reduce the impact of Data Leakage, detected as the main failure mode. Following, the implementation of PoC environment, using the internal requirement data, enabling simulation of real workflows under controlled conditions.

Purpose

The main goal is to transform the conceptual framework into a functional model, enabling preliminary assessment of its technical architecture, user interaction, and performance. The prototype will serve as a controlled environment to test the core mechanisms of use cases described in the requirements analysis.

QC2 – Prototype Feasibility Gate

This gate focuses on the PoC and its structure, to determine the architecture stability, data security, and integration feasibility. Here, the corrective actions implemented during this phase are verified and evaluated, thanks to experts reviewing.

5.5.3 Validation and Evaluation

This phase aims to systematically assess the quality, reliability, and robustness of the GenAI-based ALM prototype according to KQIs previously established and automotive safety requirements. It represents a crucial stage in transitioning from proof-of-concept (PoC) validation to a structured and data-driven evaluation, ensuring that the system meets both technical expectations and compliance standards. Functional and performance testing focusing on accuracy, processing time, consistency of generated outputs and compliance with maintainability, reliability, security and usability standards, will be performed. MCA will evaluate its quantitative metrics, namely, precision, recall and accuracy through a Ground Truth file, which contains the correct set of input-output combinations. If the model's output matches the expected one, it demonstrates reliable performance; otherwise, discrepancies indicate areas where the model's behaviour diverges from the desired outcome. A PFMEA assessment will be conducted at the end of this phase to compare the potential failure modes already defined in this thesis, combined with the severity, occurrence and detectability previously foreseen. During this phase, the focus is on the product's reliability and usability, and on the corrective actions defined in the paragraph 6.2, regarding the GenAI module and API, to reduce the severity of their failure modes, such as the Hallucinations and Data Transfer Failure, which aim to discourage the compliance to these quality standards.

Purpose

The goal is to measure the protype's performance and adherence to the quality standards defined above, through the measurement of the KQIs described in the Chapter 4 of this thesis and by applying quantitative metrics commonly used by MCA for tool comparison, such as precision, recall and accuracy.

QC3 – Validation & Risk Gate

The transition to scaling occurs only after successful functional and risk validation. This includes the evaluation of the corrective actions regarding both the Hallucinations and Data Transfer failure modes, through the comparison between the outcomes of the DFMEA and the RPNs resulted after their actual implementation. The latter evaluation, combined with the KQIs measurements, will be analysed by experts which will decide whether the product is ready for the commercial phase or not.

5.5.4 Scaling and Industrialization

The final phase of the roadmap addresses the progressive transition from prototype to deployable solution, focusing on scalability, maintainability, and commercial potential. It includes: refining the system architecture for scalability and security, defining the deployment strategy, conducting make-or-buy evaluations, performing cost–benefit and sustainability analyses and the corrective action aimed at reducing the possible Loss of Requirements Traceability failure mode. A commercialization strategy will also be developed to define the service model and target clients in the automotive sector. Quality assurance will rely on FMEA outcomes and a risk-monitoring dashboard, supported by MCA Engineering's internal quality gate system.

Purpose

The primary objective of this phase is to industrialize the GenAI-ALM integration, enabling its adoption across multiple internal teams and, subsequently, its potential offering to clients as a service or integrated product. The phase also aims to ensure operational readiness, defining governance, cost models, and continuous improvement processes for future iterations.

QC4- Deployment Readiness Gate

The last gate verifies the operational readiness of the GenAI-enhanced ALM platform, its maintainability and performance criteria compliance. It marks the final approval for internal or commercial deployment, ensuring system stability, data governance adherence and quality criteria.

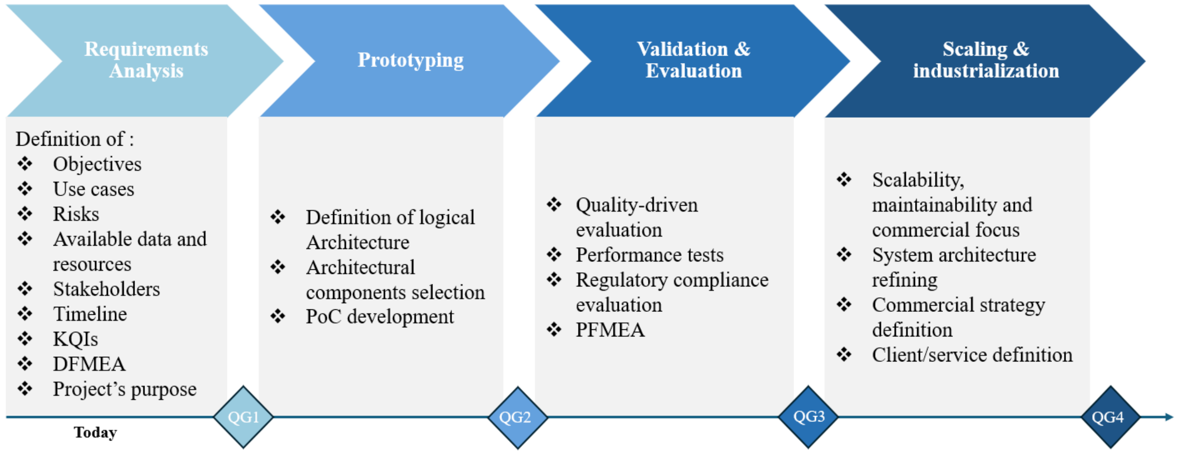

Figure 10. Project Roadmap.

6. Evaluation of Benefits, Risks, and Quality

6.1 Cost of Quality (CoQ)

Cost of Quality (CoQ) is a methodology employed by the organizations to understand, quantify and reduce the expenses associated with maintaining the quality of a product or a service. COQ comprises the expenses incurred to prevent quality issues, known as Cost of Good Quality (COGQ), and the Cost of Poor Quality (COPQ), which represents the costs incurred due to product defects, nonconformities or failures. Both categories are further divided in two distinct components. The COGQ includes Prevention Costs, associated with measures aimed at reducing or eliminating defect before they occur and the Appraisal costs, sustained to evaluate and verify the quality of the product. Conversely, COPQ comprises Internal Failure Costs, associated with defects identification before the product leaves the organization and External Failure Costs, which arise after the product reach the customer. Summing up all these components an organization can quantify the Cost of Quality of a product.

Applying this framework to a GenAI-powered ALM platform proves particularly effective in assessing the organizational investment necessary to assure the quality of such an emerging and complex technology. Although precise cost quantification remains challenging, given the immaturity of the current market and the exploratory nature of GenAI integrations, it is still possible to delineate the types of actions and corresponding expenditures associated with each component. These activities,can be outlined as follows:

Cost of Good Quality (COGQ):

1. Prevention Costs, incurred to eliminate errors and mitigate hallucinations produced by the GenAI tool, including [60] :
   - Training employees in prompt engineering and regulatory compliance (e.g., ISO 26262) and ethical AI usage.
   - Fine-tuning of GenAI models on proprietary, domain-specific databases.

- Development and implementation of trust layers and zero-retention policies to ensure data protection and accuracy.

2. Appraisal Costs: related to verifying the correctness and compliance of GenAI-generated outputs (e.g., requirements or test cases), and are [60]:
    - Design and execution of automated and manual validation routines.
    - Code reviews and regulatory cross-referencing with regulatory standards.
    - Maintenance of version-controlled documentation to ensure traceability and auditability.

Cost of Poor Quality (COPQ):

3. Internal Failure Costs: associated with GenAI errors detected before release, such as hallucinated or irrelevant requirements and semantic misinterpretation of context, which may cause rework or project delays.

4. External Failure Costs: incurred when failures are identified after deployment, such as non-compliance with automotive safety standard, customer dissatisfaction due to faulty specifications or test cases or potential product recalls, re-certifications or legal liabilities arising from undetected GenAI faults.

6.1.1 Cost-benefit relationship

As previously outlined, the integration of a Generative AI (GenAI) into an ALM platform involves both costs and benefits. Determining whether the advantages outweigh the associated costs is critically complex, as both dimensions vary significantly depending on an organization's strategic objectives, operational scale, and technological maturity. Nevertheless, the interdependence between these factors is evident, and this integration can be seen as a means for a broader transformation in enterprise management practices.

Among the most impactful benefits there is the reduction of project cycle times, primarily achieved through the automation of resource-intensive tasks such as test case generation, requirements review, and code documentation. Early industry pilots suggest that GenAI can reduce development times by 30-50%, translating into significant operational cost savings and efficiency gains. Furthermore, the implementation of natural language processing (NLP) techniques, as the translation of informal business language into formal technical specifications, enhances the consistency and quality of requirements. This improvement also supports compliance with industry standards, such as ISO 26262 and ASPICE, which is especially critical in the automotive sector [50]. GenAI tools also promote cross-functional collaboration through intelligent user interfaces that enable the access to technical information even to non-specialized professionals, by bridging communication gaps between business and technical teams. Additionally, GenAI could improve significantly scalability, enabling the expansion of development initiatives without necessitating a proportional increase in headcount. Some studies report productivity gains of up to 44% when GenAI is integrated with quality assurance mechanisms, and time savings of 40% or more in tasks involving code writing, testing, and documentation [37], [60].

However, these benefits lead to substantial initial investments. Organizations opting for open-source solutions, for instance, must acquire powerful computational hardware capable of supporting large language models (LLMs), a prohibitive requirement for small to mid-sized firms. Equally important is the need for employee training in prompt engineering, regulatory compliance, and responsible AI usage to fully leverage GenAI's capabilities. Moreover, integration-related costs, including API development, data infrastructure, cybersecurity measures, and continuous testing protocols, must be considered. These expenses are both part of the costs of quality (CoQ) and of the initial investments to obtain a product that ensures compliance, performance, and reliability in enterprise settings [37], [60].

Even if the economic implications of GenAI integration vary case by case, the strategic value it offers in terms of time efficiency, quality assurance, collaboration, and scalability makes it a powerful innovation. Organizations that invest in both infrastructure and human capital are better positioned to reap long-term returns, not only through cost savings but also through enhanced agility and competitive differentiation.

Table 3 - Quality cost- benefit relationship

Benefits		Costs	
Category	Estimated impact	Category	Estimated impact
Development Time Reduction	30–50% cycle time reduction	Hardware investment	High upfront capital expenditure
Quality Improvement	Improved compliance, reduced rework	Employee training	Medium to high depending on team size
Cross-Team Collaboration	Faster feedback loops, fewer miscommunications	API & Infrastructure Integration	Moderate to high, depends on platform architecture
Scalability	Up to 44% productivity gain	Security & Compliance testing	Continuous cost due to regulatory evolution

6.2 Risk Analysis and Mitigation

6.2.1 FMECA for GenAI systems

The Failure modes, effects and criticality analysis (FMECA) is a methodology to identify and analyse all potential failure modes within a system, assess their effects on overall functionality, and determine appropriate strategies to prevent or mitigate their impact. This approach plays a crucial during both the early stages of design, and the deployment process, as it supports the selection of alternative which maximize reliability and safety. In fact, it could be distinguished in DFMEA during the Design phase and PFMEA during the Process stage. Before performing an FMEA, define the system boundaries, the product objectives, and the operational environment is necessary. Developing a Functional Block Diagram (FBD) assists the visualization of the system's key functions and the definition of the appropriate level of analysis. Subsequently, the preparation of the FMEA worksheet enables the identification of potential failure modes and provides a structured basis for assessing corrective and preventive actions. Within the worksheet, each component is examined in relation to its functions to determine whether potential failures could produce unacceptable effects on other components or on the overall system, in terms of functionality degradation. Each failure mode is then evaluated according to three key parameters: Severity (S), representing the gravity of the failure's consequences; Occurrence (O), reflecting the likelihood of the failure mode and its cause; and Detection (D), indicating the probability that the failure will remain undetected. These three are quantified using numerical scale ranging from 0 to 10, this project used the one suggested by the Automotive Industry Action Group (AIAG 2019) illustrated in Figure 2. Higher severity and occurrence value correspond to more severe or frequent failures, while a higher detection value denotes a lower likelihood of detection. The product of these three indicators is the Risk Priority Number (RPN), which reveals the criticality of each failure mode-cause combination. However, since the RPN alone may not provide a comprehensive assessment for risk prioritisation, during the team review phase, additional tools such as the Pareto chart and Risk matrix are employed to support a more informed assessment and guide the selection of the most appropriate corrective actions [68].

Rating	Severity	Occurr.(O)	Detenction
10	May endanger (machine or assembly) operator, without warning.	Very High (P > 10%)	No current process control; failure mode and/or cause cannot be detected or analyzed.
9	May endanger (machine or assembly) operator, with warning.	High (P > 5%)	Failure mode and/or cause is not easily detected (e.g., random audits).
8	100% of product may have to be scrapped; the production line may be shut down.	High (P ≈ 2%)	Failure mode detection post-processing by operator through visual/tactile/audible means.
7	Product may have to be sorted and a portion scrapped; production line still operational.	Moderate (P > 1%)	Detection in-station by operator (visual/tactile) or post-processing via attribute gauging (e.g., go/no-go, manual torque check, etc.).
6	100% of the product may have to be reworked off-line and then accepted/rejected.	Moderate (P ≈ 0.2%)	Detection post-processing by operator or in-station using variable or attribute gauging.
5	A portion of the product may have to be reworked off-line and then accepted/rejected.	Moderate (P ≈ 0.05%)	Detection in-station by operator or automated controls (light, buzzer). Setup and first-piece checks also used.
4	100% of product may have to be reworked in-station before being processed.	Moderate (P ≈ 0.01%)	Detection post-processing by automated controls; lockout of defective parts.
3	A portion of product may have to be reworked in-station before being processed.	Low (P ≈ 0.001%)	Detection in-station by automated controls; lockout of defective parts in-station.
2	Slight inconvenience to process, operation, or operator.	Low (P ≤ 0.0001%)	Detection in-station by automated controls that detect errors and prevent operation.
1	No discernible effect.	Very Low (P ≈ 0%)	Prevention by design (fixture, machine, or part) — error-proofed design prevents issue.

Figure 11. Severity/occurrence/detection evaluation criteria by AIAG 2019.

This thesis performed a DFMEA of a GenAI-enhanced ALM platform to identify its potential failure modes and assess the severity of their consequences. The system analysed comprises four main architectural components: the ALM Core, the GenAI module, the secure database, and the API layer. Its primary mission is to enhance Requirements Management by automating critical tasks such as test case generation, ambiguity detection, and documentation production, while ensuring compliance with automotive safety and quality standards. The operational environment is expected to be compliant with the strict regulations and able to facilitate teams' collaboration and handle sensitive data. The system must deal with challenges such as protecting data, managing high computational loads and ensuring compatibility with other systems, in both cloud and company servers deployment scenarios.

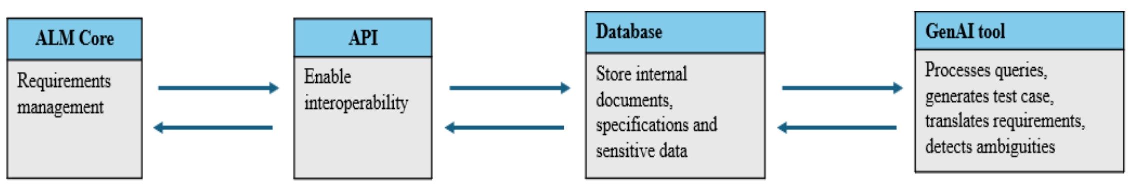

Figure 12. Functional Block Diagram of the GenAI-ALM system.

The research focused on foreseeing and identifying the failure modes that effects the most the four quality attributes defined by the ISO/IEC 25010 standard, namely: maintainability, security, usability, and reliability. The methodological approach guided by qualitative analysis and MCA experts, involved the selection and evaluation of the predominant failure mode for each system component considering its potential to compromise these critical system qualities. For the ALM Corse, the Loss of Requirements Traceability was identified as the main failure mode, since the non-compliance regulatory consequences effects negatively system reliability, usability and maintainability. Within the secure database, the Data Leakage emerged as the most severe failure mode (RPN = 560), threatens data confidentiality and integrity, thus posing a major risk to security, a primary concern in the automotive domain. The major criticality detected in API layer, was the Data Transfer Failure, which may lead to the corruption of generated content due to inadequate schema validation. This failure directly compromises the system's usability by disrupting communication between components. Finally, the GenAI was associated with the failure mode of Hallucination, defined as the generation of no-compliant or misleading requirements. With an RPN of 504, this represents the second most critical risk and undermines the platform's reliability and maintainability, particularly in scenarios where fine-tuning is insufficient.

Unique ref.	Function	Failure modes	Effect	S	Failure Cause(s)	O	Detection system	D	RPN
ALM Core	Requirements management	Loss of requirements traceability	Regulatory non-compliance	9	Lack of automated trace propagation on change requests	6	Manual review of traceability matrices	6	324
Database	Sensitive Data storage	Data Leakage	Confidentiality and contract violation	10	Weak access control	7	Anomaly detection	8	560
API	Component comunication	Data transfer failure	Corruption of generated content	8	Lack of schema validation	5	Response validation	6	240
GenAI	Requirements generation	Hallucination	Non-compliant or misleading requirements	9	Insufficient fine-tuning	7	Human-in-the-loop validation	8	504

Figure 13. FMEA worksheet of GenAI-enhanced ALM platform.

To support the RPNs outcomes, a Risk matrix, constructed using severity and occurrence values, and a Pareto chart, were developed.

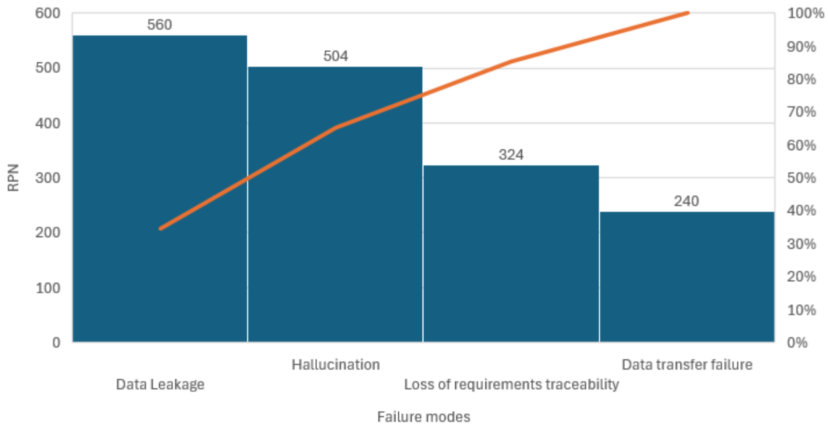

Figure 14. Pareto Chart.

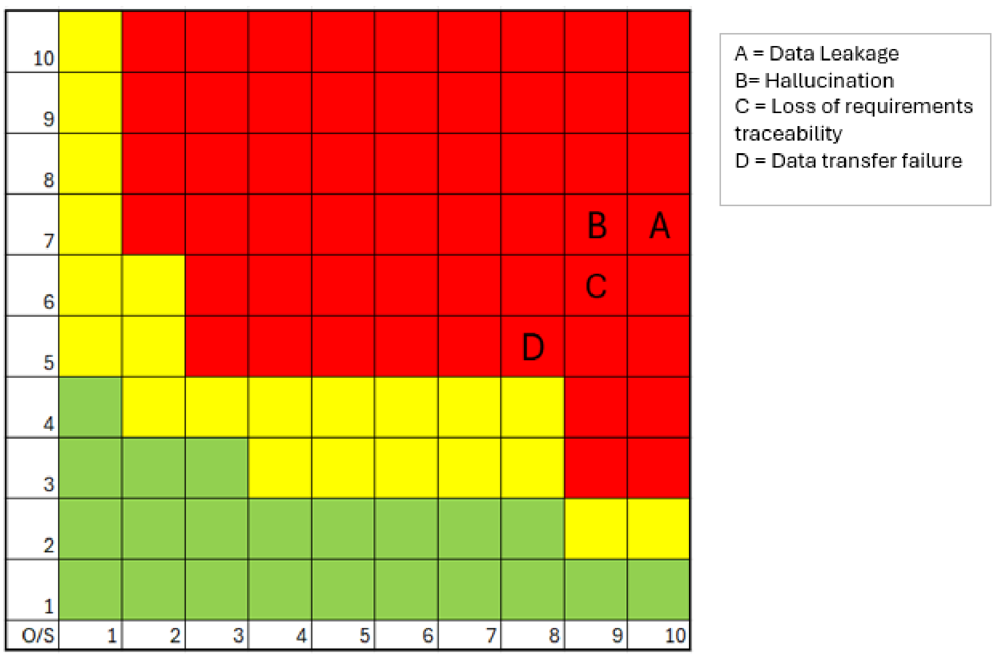

Figure 15. Risk Matrix.

The outcomes revealed by the three tools above, show as the first two major failure modes, the Data Leakage and the Hallucinations, followed by the Loss of requirements traceability and Data transfer failure. Corrective actions were defined with the objective of reduce their criticality during the subsequent product development phase.

The corrective actions definition starts with the most critical failure mode, which is the Data Leakage. The following measures identified, aimed to decrease its occurrence and increases its detectability, will be employed and evaluated, respectively, during the prototyping phase and its quality date (QC2). The corrective actions are:

- *Application of zero-retention policies,* in cases involving the use of commercial GenAI systems, whereby data is deleted immediately after task completion.
- *Regular audits of data governance*, with the purpose of ensuring continuous compliance with security standards and detecting potential liabilities in data handling processes.

The measure detected to reduce the severity and enhance the detectability of Hallucination, in the GenAI module, are the following, and:

- *Human-in-the-loop implementation,* ensuring that generated content is reviewed and confirmed before utilization.
- *Regular fine-tuning* of the model using domain-specific and continuously updated datasets.

They will be applied and assessed during the Validation & Evaluation phases and its quality gate (QG3), since they would mitigate the criticality of the failure mode during the testing.

Since the Loss of requirements traceability of the ALM Core, becomes more critical during the operational adoption, the following corrective actions, will be conducted in the Scaling and industrialization phase and evaluated in its relative gate (QG4). They aim to reduce the occurrence and of increase the overall control of this failure mode:

- *Enhanced team training* on traceability workflows and tools, improving consistency in change management procedures.
- *Periodic traceability audits*, to ensure the integrity of requirement linkages across the lifecycle.

The Data Transfer Failure in the API component was analysed last due to its lower criticality ranking. The main objectives, as for the ALM one, were to reduce the occurrence and enhance the control, through the resulting methods:

- *Validation of input/output schemas*, ensuring compatibility and consistency of data structures during transfer.
- *Increasing end-to-end testing of the API interface* before the deployment stage, to identify integration issues early and ensure robust system communication.

These actions are comprised in the Validation and Evaluation phase and will be assessed in the respective quality gate (QG3), ensuring that the modules communication is stable and without errors.

Table 4 - Corrective action and Quality Gate connection.

Failure mode	Corrective action	Quality Gate
Data Leakage	Application of zero retention policies	2
	Regular audits of data governance	
Hallucination	Human-in-the-loop implementation,	3
	Regular fine-tuning	
Loss of requirements traceability	Enhanced team training	4
	Periodic traceability audits,	
Data transfer failure	Validation of input/output schemas	3
	Increasing end-to-end testing of the API interface	

Conclusions

This thesis was developed in collaboration with MCA Engineering S.r.l, with the purpose of proposing a GenAI-enhanced ALM platform, able to address the challenges, process complexity and safety constraints related to the Requirements Management in the Automotive sector. The research aimed to define a quality-driven framework which could thin the academic gap in this field and provide a roadmap for the implementation of this system in an immature market.

The work presented a conceptual framework and a candidate architecture for a GenAI-enhanced ALM platform, focusing on the quality compliance with the maintainability, reliability, security and usability standards, defined by the ISO/IEC 25010. Moreover, KQIs, related to the latter features and the relative sub-features, were determined from the mathematical functions described in ISO/IEC 25023. The project roadmap was proposed with the objective of supporting the conceptualization and future development of GenAI-enhanced ALM platform. The project roadmap divides the process in four stages, starting from the Requirements Analysis, followed by the Prototyping phase, then the Validation and Evaluation, ending with the Scaling and industrialization. Each of these phases is associated with a Quality Gate, essential to reflect the quality-driven approach and to validate the accuracy and compliance of each stage, before proceeding to the following one. The thesis includes a Risk analysis, a Cost of Quality definition and a Cost-Benefit relationship, to identify the criticality issues in undertaking this project. The corrective actions determined during the DFMEA will be deployed during the following phases of the project and evaluated in the respective gate, to quantify their effectiveness.

The main contribution lies in the establishment of a structured, measurable, and quality-oriented approach for the integration of GenAI into Application Lifecycle Management platform. The model contributes academically to the still immature field of GenAI- ALM integration. The activities conducted in this thesis enabled MCA Engineering to accomplish the first stage of the roadmap, providing a complete Requirement Analysis framework, establishing the conceptual and methodological foundations for the subsequent prototyping, validation and industrialization steps.

Since the project is still in its early stages, the proposed framework has not yet undergone extensive experimental validation. Consequently, the absence of publicly available industrial implementations of GenAI-ALM integration prevented a detailed and quantitative cost-benefit or risk-return analysis, leaving these evaluations in a qualitative stage. Therefore, future work will focus on implementing the proof-of-concept (PoC) prototype, executing quantitative quality assessments based on ISO/IEC 25023 metrics, and evaluating the effectiveness of the corrective actions defined in the FMEA analysis. Furthermore, the industrial deployment of the system will enable the progressive refinement of the model and the assessment of its economic impact.

The thesis provided aims to bridge the academic research gap and enterprise practice, regarding the GenAI-ALM integration as a new tool to enhance the Requirements Management process in the automotive industry. The adoption of a rigorous Quality Engineering approach outlined how should be a responsible approach of Generative AI in complex industrial environments, such as the automotive one. It contributes to building safer, more reliable systems, able to achieve the compliance regulations and quality standards.

From a broader perspective, this work provides a structured foundation for future research and industrial applications. It demonstrates that the integration of Generative Artificial Intelligence into enterprise software platforms, fostered by the Quality Engineering principles, can be a replicable and effective model for developing secure, traceable, and efficient digital ecosystems. This qualitative focus not only supports the reliability and sustainability of the project itself but also ensures that innovation remains aligned with quality assurance and regulatory compliance.

References

[1] R. Gupta, K. Nair, M. Mishra, B. Ibrahim, and S. Bhardwaj, “Adoption and impacts of generative artificial intelligence: Theoretical underpinnings and research agenda,” *International Journal of Information Management Data Insights*, vol. 4, no. 1, Apr. 2024, doi: 10.1016/j.jjimei.2024.100232.

[2] R. Merritt, “What Is a Transformer Model? | NVIDIA Blogs.” Accessed: Nov. 03, 2025. [Online]. Available: https://blogs.nvidia.com/blog/what-is-a-transformer-model/

[3] “What is Generative AI and How Does it Work? | NVIDIA Glossary.” Accessed: Sep. 27, 2025. [Online]. Available: https://www.nvidia.com/en-us/glossary/generative-ai/

[4] “What is Synthetic Data Generation (SDG)? | NVIDIA Glossary.” Accessed: Sep. 27, 2025. [Online]. Available: https://www.nvidia.com/en-us/glossary/synthetic-data-generation/

[5] A. Singla, A. Sukharevsky, L. Yee, M. Chui, and B. Hall, “The state of AI How organizations are rewiring to capture value,” 2025.

[6] “Get AI Ready: Action plan for IT Leaders | Gartner.” Accessed: Sep. 27, 2025. [Online]. Available: https://www.gartner.com/en/information-technology/topics/ai-readiness

[7] “Deloitte’s State of Generative AI in the Enterprise Quarter four report,” 2025. Accessed: Sep. 28, 2025. [Online]. Available: https://www.deloitte.com/content/dam/assets-zone3/us/en/docs/campaigns/2025/us-state-of-gen-ai-2024-q4.pdf

[8] “A new era of generative AI for everyone,” May 2023. Accessed: Sep. 28, 2025. [Online]. Available: https://www.accenture.com/content/dam/accenture/final/accenture-com/document/Accenture-A-New-Era-of-Generative-AI-for-Everyone.pdf

[9] “Generative Tomorrow The Future Unveiled, Starring GenAI,” Milan, May 2025. Accessed: Sep. 28, 2025. [Online]. Available: https://www.deloitte.com/content/dam/assets-zone2/it/it/docs/events/2025/generative-tomorrow-ita.pdf

[10] N. Research Institute, “Sovereign AI has spurred wave of new AI strategies,” 2025. [Online]. Available: https://commission.europa.eu/topics/

[11] "Introducing ChatGPT Enterprise | OpenAI." Accessed: Sep. 28, 2025. [Online]. Available: https://openai.com/index/introducing-chatgpt-enterprise/?

[12] "OpenAI's ChatGPT Enterprise: Pricing, Benefits, and Security." Accessed: Sep. 28, 2025. [Online]. Available: https://meetcody.ai/blog/open-ai-chatgpt-enterprise-pricing-benefits-security/?

[13] "PwC is accelerating adoption of AI with ChatGPT Enterprise in US and UK and with clients : PwC." Accessed: Sep. 28, 2025. [Online]. Available: https://www.pwc.com/us/en/about-us/newsroom/press-releases/pwc-us-uk-accelerating-ai-chatgpt-enterprise-adoption.html

[14] "Accelerating the development of life-saving treatments | OpenAI." Accessed: Sep. 28, 2025. [Online]. Available: https://openai.com/index/moderna/

[15] "Meet Breeze — HubSpot's AI tools that make impossible growth impossibly easy." Accessed: Sep. 28, 2025. [Online]. Available: https://www.hubspot.com/products/artificial-intelligence?utm_id=687188076456&utm_term=sales_hubspot%20ai_EN&utm_campaign=Sales_Portals_EN_NAM_NAM_AIGeneral-Brand_e_c_campaignid20924742702_agid158136316992_google&utm_source=google&utm_medium=paid&utm_content=_&hsa_acc=9694350438&hsa_cam=20924742702&hsa_grp=158136316992&hsa_ad=687188076456&hsa_src=g&hsa_tgt=kwd-386212193500&hsa_kw=hubspot%20ai&hsa_mt=e&hsa_net=adwords&hsa_ver=3&gad_source=1&gclid=CjwKCAiAzvC9BhADEiwAEhtlN_ZPnBSHMf7pO6qwitTmD-axSAgJ0VVRq6kn9am_3e9rCg4Vb_NtjxoCdlAQAvD_BwE#below-header-breeze_wf_header_splash

[16] "Timely, automatic time tracking for consultancies, agencies and SaaS businesses." Accessed: Sep. 28, 2025. [Online]. Available: https://www.timely.com/

[17] "Generative AI agents: The future of enterprise solutions | ZS." Accessed: Sep. 28, 2025. [Online]. Available: https://www.zs.com/insights/generative-ai-agents-for-the-modern-enterprise

[18] S. Kim, J. Ha, H. Lee, S. Park, and S. Cho, "Human-guided collective LLM intelligence for strategic planning via two-stage information retrieval," *Inf Process Manag*, vol. 63, no. 1, Jan. 2026, doi: 10.1016/j.ipm.2025.104288.

[19] “Three New AI Breakthroughs Shaping 2026: AI Trends | Deloitte US.” Accessed: Sep. 28, 2025. [Online]. Available: https://www.deloitte.com/us/en/services/consulting/blogs/new-ai-breakthroughs-ai-trends.html?

[20] “The Dawn of Humanoid Robots and Physical AI - Leo Wealth.” Accessed: Sep. 28, 2025. [Online]. Available: https://leowealth.com/insights/the-dawn-of-humanoid-robots-and-physical-ai/?

[21] Singla Alex *et al.*, “What is multimodal AI?,” Jun. 2025. Accessed: Sep. 28, 2025. [Online]. Available: https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-multimodal-ai

[22] Rishi Bommasani, Drew A. Hudson, and Ehsan Adeli, “On the Opportunities and Risks of Foundation models,” *ArXiv*, 2021, Accessed: Sep. 28, 2025. [Online]. Available: https://crfm.stanford.edu/assets/report.pdf

[23] “What Are Foundation Models? | IBM.” Accessed: Sep. 28, 2025. [Online]. Available: https://www.ibm.com/think/topics/foundation-models?

[24] “What Are Foundation Models? | NVIDIA Blogs.” Accessed: Sep. 28, 2025. [Online]. Available: https://blogs.nvidia.com/blog/what-are-foundation-models/

[25] “What Is a Pretrained AI Model? | NVIDIA Blog.” Accessed: Sep. 28, 2025. [Online]. Available: https://blogs.nvidia.com/blog/what-is-a-pretrained-ai-model/

[26] “What Is A Pretrained Model? | IBM.” Accessed: Sep. 28, 2025. [Online]. Available: https://www.ibm.com/think/topics/pretrained-model

[27] “Open Sourcing BERT: State-of-the-Art Pre-training for Natural Language Processin.” Accessed: Sep. 28, 2025. [Online]. Available: https://research.google/blog/open-sourcing-bert-state-of-the-art-pre-training-for-natural-language-processing/

[28] L. Stappen, J. Dillmann, S. Striegel, H.-J. Vögel, N. Flores-Herr, and B. W. Schuller, “Integrating Generative Artificial Intelligence in Intelligent Vehicle Systems,” May 2023, [Online]. Available: http://arxiv.org/abs/2305.17137

[29] “Open Source vs Commercial GPT: AI Model Comparison.” Accessed: Sep. 28, 2025. [Online]. Available: https://www.mycustomai.io/blog/ai-landscape-open-source-vs-commercial-ai-guide?

[30] "Commercial vs. Open-Source GenAI: What to Choose? - Spiceworks." Accessed: Sep. 28, 2025. [Online]. Available: https://www.spiceworks.com/tech/artificial-intelligence/guest-article/commercial-vs-open-source-genai/

[31] "MBUX: sistema multimediale con AI integrata | Mercedes-Benz." Accessed: Sep. 28, 2025. [Online]. Available: https://www.mercedes-benz.it/passengercars/technology/mbux.html?%3Fcsref=_sm%3Ain_twr_pc

[32] "Audi enhances voice control in current and future models with ChatGPT | Audi MediaCenter." Accessed: Sep. 28, 2025. [Online]. Available: https://www.audi-mediacenter.com/en/press-releases/audi-enhances-voice-control-in-current-and-future-models-with-chatgpt-16068

[33] "bigscience/bloom · Hugging Face." Accessed: Nov. 03, 2025. [Online]. Available: https://huggingface.co/bigscience/bloom?

[34] "Introducing the Technology Innovation Institute's Falcon 3 Making Advanced AI accessible and Available to Everyone, Everywhere." Accessed: Nov. 03, 2025. [Online]. Available: https://falconllm.tii.ae/falcon-180b.html

[35] Hermansen Anna and Osborne Cailean, "The Economic and Workforce Impacts of Open Source AI THE ECONOMIC AND WORKFORCE IMPACTS OF OPEN SOURCE AI | 3," May 2025. Accessed: Sep. 28, 2025. [Online]. Available: https://www.linuxfoundation.org/hubfs/Research%20Reports/lfr_marketimpacts25_052725a.pdf?hsLang=en

[36] "How open source AI solutions are reshaping business | McKinsey." Accessed: Sep. 28, 2025. [Online]. Available: https://www.mckinsey.com/capabilities/quantumblack/our-insights/open-source-technology-in-the-age-of-ai?

[37] Hepp Dominik, Harrysson Martin, Maslaniec Lukasz, Wozniak Mateusz, and Amroudi Micheal, "From engines to algorithms: Gen AI in automotive software developmentdevelopment," *McKinsey & Company*, Accessed: Sep. 28, 2025. [Online]. Available: https://www.mckinsey.com/features/mckinsey-center-for-future-mobility/our-insights/from-engines-to-algorithms-gen-ai-in-automotive-software-development

[38] "Calculating the Cost of Generative AI — ITRex." Accessed: Sep. 28, 2025. [Online]. Available: https://itrexgroup.com/blog/calculating-the-cost-of-generative-ai/?

[39] "Fine-tuning now available for GPT-4o | OpenAI." Accessed: Sep. 28, 2025. [Online]. Available: https://openai.com/index/gpt-4o-fine-tuning/

[40] "Generative AI in automotive industry: Use cases, solution and implementation." Accessed: Sep. 28, 2025. [Online]. Available: https://www.leewayhertz.com/generative-ai-in-automotive-industry/

[41] "Generative AI in Automotive | IBM." Accessed: Sep. 28, 2025. [Online]. Available: https://www.ibm.com/think/topics/generative-ai-automotive#footnotes5

[42] N. Petrovic *et al.*, "Survey of GenAI for Automotive Software Development: From Requirements to Executable Code," Jul. 2025, [Online]. Available: http://arxiv.org/abs/2507.15025

[43] "Volkswagen to Invest $1.2 Billion in Artificial Intelligence by 2030 - WSJ." Accessed: Sep. 28, 2025. [Online]. Available: https://www.wsj.com/tech/ai/volkswagen-to-invest-1-2-billion-in-artificial-intelligence-by-2030-e41b2021

[44] Venus Andreas, Poltronieri Giulietta, de Jond Marc, Kellner Martin, Blos Alina, and Steiner Fabian, "The rise of edge AI in automotive." Accessed: Sep. 28, 2025. [Online]. Available: https://www.mckinsey.com/industries/se%20miconductors/our-insights/the-rise-of-edge-ai-in-automotive

[45] "Google Gemini sta per arrivare sulla tua Volvo con Google integrato." Accessed: Sep. 28, 2025. [Online]. Available: https://www.volvocars.com/it/news/technology/google-gemini-is-coming-to-your-volvo-with-google-built-in/

[46] "L'intelligenza artificiale come motore per l'ottimizzazione del controllo qualità." Accessed: Sep. 28, 2025. [Online]. Available: https://www.press.bmwgroup.com/italy/article/detail/T0449908IT/l-intelligenza-artificiale-come-motore-per-l-ottimizzazione-del-controllo-qualit%C3%A0

[47] "What Is Requirements Management? | IBM." Accessed: Sep. 28, 2025. [Online]. Available: https://www.ibm.com/think/topics/what-is-requirements-management

[48] "Requirements Management for Automotive Development Guide - Visure Solutions." Accessed: Sep. 28, 2025. [Online]. Available:

https://visuresolutions.com/automotive/requirements-management/#elementor-toc__heading-anchor-7

[49] "Requirements Management in Automotive Engineering 2025." Accessed: Sep. 28, 2025. [Online]. Available: https://www.enco-software.com/requirements-management-in-automotive-engineering/?

[50] "Reshaping of ALM using Generative AI - Kovair Blog." Accessed: Sep. 28, 2025. [Online]. Available: https://www.kovair.com/blogs/generative-ai-the-missing-piece-in-end-to-end-application-lifecycle-management-alm/?

[51] *BSI Standards Publication Systems and software engineering-Systems and software Quality Requirements and Evaluation (SQuaRE)-Product quality model*. 2023.

[52] *Systems and software engineering-Systems and software Quality Requirements and Evaluation (SQuaRE)-Measurement of system and software product quality BSI Standards Publication*. 2016.

[53] H. Cheng *et al.*, "Generative AI for Requirements Engineering: A Systematic Literature Review," 2025.

[54] Y. Uygun and V. Momodu, "Local large language models to simplify requirement engineering documents in the automotive industry," *Prod Manuf Res*, vol. 12, no. 1, 2024, doi: 10.1080/21693277.2024.2375296.

[55] "Volkswagen enhances engineering efficiency with Microsoft Copilot and PTC | Microsoft Customer Stories." Accessed: Oct. 10, 2025. [Online]. Available: https://www.microsoft.com/en/customers/story/24120-volkswagen-microsoft-copilot?

[56] "Transforming ALM with HCLTech ALMate on AWS | HCLTech." Accessed: Oct. 10, 2025. [Online]. Available: https://www.hcltech.com/blogs/revolutionizing-application-lifecycle-management-hcltech-almate-for-siemens-polarion-powered-by-aws

[57] M. Arora and A. Kumar, "An Empirical Study on Make-or-buy Decision Making," *International Journal of Education and Management Engineering*, vol. 12, no. 1, pp. 19–28, Feb. 2022, doi: 10.5815/ijeme.2022.01.03.

[58] "16 Best Requirements Management Software for 2025 | Pros and Cons - Visure Solutions." Accessed: Oct. 13, 2025. [Online]. Available: https://visuresolutions.com/alm-guide/top-requirements-management-software-tools/

[59] "10+ Best Requirements Engineering Tools & Softwares for 2025 - Visure Solutions." Accessed: Oct. 13, 2025. [Online]. Available: https://visuresolutions.com/blog/best-requirements-engineering-tools-and-softwares/

[60] D. Safaei and K. Haki, "GENERATIVE ARTIFICIAL INTELLIGENCE FOR ENTERPRISE PLATFORMS," Cyprus, Jun. 2024.

[61] "Codebeamer—ALM Solution | PTC." Accessed: Oct. 20, 2025. [Online]. Available: https://www.ptc.com/en/products/codebeamer

[62] "OSLC and Jazz for Service Management - IBM Documentation." Accessed: Oct. 20, 2025. [Online]. Available: https://www.ibm.com/docs/en/jfsm/1.1.2.1?topic=overview-oslc-jazz-service-management

[63] "What is RAG (Retrieval Augmented Generation)? | IBM." Accessed: Oct. 21, 2025. [Online]. Available: https://www.ibm.com/think/topics/retrieval-augmented-generation

[64] "What Is an API (Application Programming Interface)? | IBM." Accessed: Oct. 21, 2025. [Online]. Available: https://www.ibm.com/think/topics/api

[65] "Integrating data by using the OSLC REST API - IBM Documentation." Accessed: Oct. 21, 2025. [Online]. Available: https://www.ibm.com/docs/en/masv-and-l/maximo-ref/cd?topic=applications-integrating-data-by-using-oslc-rest-api&utm_source=chatgpt.com

[66] "REST API (v1)." Accessed: Oct. 21, 2025. [Online]. Available: https://support.ptc.com/help/codebeamer/r2.2/en/index.html#page/codebeamer/developers_guide/117612.html

[67] "LLM Retrieval-Augmented Generation (RAG) with OpenVINO™ and LangChain 1. Executive Summary."

[68] Franceschini Fiorenzo, Galetto Maurizio, and Maisano Domenico, *Designing Performance Measurement Systems Theory and Practice of Key Performance Indicators*. Springer International Publishing, 2019. doi: https://doi.org/10.1007/978-3-030-01192-5.