

Article

Not peer-reviewed version

A Generative AI-Based Framework for Proactive Quality Assurance and Auditing

[Galina Ilieva](#)*, [Tania Yankova](#), [Vera Hadzhieva](#), [Yuliy Iliev](#)

Posted Date: 8 January 2026

doi: 10.20944/preprints202601.0579.v1

Keywords: generative AI (AI); quality management (QM); quality assurance (QA); artificial intelligence (AI); large language models (LLMs); automated quality control; digital quality audits; quality standard compliance; smart manufacturing; human-in-the-loop systems



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

A Generative AI-Based Framework for Proactive Quality Assurance and Auditing

Galina Ilieva ^{1,*}, Tania Yankova ¹, Vera Hadzhieva ¹ and Yuliy Iliev ²

¹ Department of Management and Quantitative Methods in Economics, University of Plovdiv Paisii Hilendarski, 4000 Plovdiv, Bulgaria

² Teletek Electronics, 1407 Sofia, Bulgaria

* Correspondence: galili@uni-plovdiv.bg

Abstract: Generative Artificial Intelligence (AI) is transforming quality management (QM) and auditing by expanding automation, supporting data-driven decisions, and enabling more personalized stakeholder interaction. However, its adoption also raises concerns related to system robustness, operational resilience, and regulatory compliance, including potential deviations from Critical-to-Quality (CTQ) requirements, gaps in traceability, and misalignment with established quality standards. This paper proposes a structured conceptual framework for proactive, generative AI-enabled QM and auditing, organized into three functional domains: supplier performance, in-process control, and post-market feedback. The framework shows how generative AI can: 1) strengthen supplier oversight via automated documentation and early risk identification; 2) improve in-process control through real-time anomaly detection and Statistical Process Control (SPC)-based triage; and 3) enhance post-market surveillance using predictive analytics for warranty clustering and prioritized Corrective and Preventive Action (CAPA) preparation. To ensure compliance and auditability, the framework incorporates policy-based constraints, human-in-the-loop checkpoints, and end-to-end digital traceability. Verification was performed through a proof-of-concept case study spanning discrete manufacturing and process-based production environments, comparing a conventional quality workflow with a generative AI-augmented alternative. Expert assessment indicated that the generative AI-assisted workflow achieved better performance on key criteria, including documentation completeness, defect detection, process stability, governance and time efficiency. The obtained results suggest that the proposed framework can support a shift from reactive quality control towards predictive and preventive improvement while preserving alignment with quality standards and organizational quality objectives.

Keywords: generative AI (AI); quality management (QM); quality assurance (QA); artificial intelligence (AI); large language models (LLMs); automated quality control; digital quality audits; quality standard compliance; smart manufacturing; human-in-the-loop systems

1. Introduction

Digital transformation has reshaped how organizations design quality systems and ensure product and process conformity. Across industries, the widespread deployment of connected platforms for monitoring, analytics, documentation, and auditing has improved visibility but also exposed persistent weaknesses in data interoperability, end-to-end traceability, and standards-driven compliance within quality management (QM) and quality assurance (QA) systems [1,2]. These challenges are further intensified by the rapid incorporation of artificial intelligence (AI), which is changing how assurance activities are planned, executed, and evidenced in everyday quality operations [3].

Recent advances in generative AI and large language models (LLMs) [4,5] have introduced systems that can interpret user intent, interact in natural language, and generate structured, task-specific content. In manufacturing, these capabilities extend beyond conversational support to

practical contributions across quality workflows – for example, drafting and updating control plans, supporting supplier oversight, assisting in-process decisions, and enabling post-market analysis. By accelerating documentation, improving access to quality knowledge, and synthesizing complex evidence, generative AI can strengthen responsiveness, support predictive analytics, and facilitate audit-ready reporting, helping organizations move from reactive control toward more continuous and evidence-based assurance.

At the same time, generative AI introduces risks for QM/QA and auditing, particularly related to decision accuracy, explainability, traceability, and regulatory conformity [6]. In response, governance frameworks and standards are emerging. The NIST AI Risk Management Framework emphasizes risk management across the AI lifecycle, including data provenance, monitoring, and human oversight [7]. ISO/IEC AI management system standards likewise focus on organizational controls for responsible development and deployment [8]. In parallel, the EU AI Act establishes a risk-based regulatory regime with requirements for technical documentation, transparency, and post-market monitoring – requirements that closely overlap with QA and audit expectations in manufacturing settings [9]. Together, these developments reinforce the need for proactive, evidence-centric mechanisms that keep AI-enabled quality processes auditable and compliant.

Given both accelerating adoption of generative AI tools in manufacturing and rising regulatory scrutiny, QM and assurance models must remain robust, reliable, and accountable in AI-augmented production environments [6]. However, despite growing deployment of generative AI on the shop-floor and in quality offices, many existing frameworks provide limited standards-aligned guidance on how to integrate these tools while preserving auditability, traceability, and clear responsibility for decisions [10,11].

This study addresses that gap by proposing and validating a structured framework for integrating generative AI across three core QM functions: supplier performance, in-process control, and post-market feedback. The framework defines roles, responsibilities, and safeguards for each function, specifying where generative AI can assist with drafting or interpretation (e.g., control plan updates, anomaly explanations, complaint clustering) and where human review and approval remain essential. The framework is evaluated through a manufacturing case study from the electronics industry, comparing conventional (human-only) workflows with generative AI-assisted counterparts.

The main objectives of this study are:

- To identify the needs, expectations, challenges, and perspectives of key stakeholders – quality engineers, production managers, suppliers, auditors/regulators, and customers – regarding the integration of generative AI into QA and auditing activities.
- To propose practical measures and guidelines for the responsible use of generative AI in regulated QM contexts, including human-in-the-loop approvals, data governance, traceability, and alignment with recognized quality standards.
- To design a conceptual framework that enables systematic, function-sensitive integration of generative AI across supplier performance, in-process control, and post-market feedback, structured along the Plan–Execute–Improve cycle.
- To validate the proposed framework in practice across the three core QM functions using process-quality thresholds and AI-governance controls.
- To provide role-specific recommendations for different stakeholder groups on the adoption, governance, and oversight of generative AI in QA and auditing.

The primary contribution of this study is the development and validation of a generative AI-enabled QA framework that supports end-to-end quality activities while emphasizing transparency, accountability, and auditability. By addressing risks, such as over-reliance on automation, limited explainability, and potential misuse, the framework supports a transition toward more predictive and preventive quality assurance aligned with evolving regulatory requirements and stakeholder expectations.

The remainder of the paper is organized as follows. Section 2 reviews advances in generative AI relevant to QM and auditing, with emphasis on LLMs and multimodal systems for industrial use. Section 3 surveys current research on generative AI applications, highlighting benefits, limitations, and standardization trends across supplier quality management, in-process quality control, and post-market surveillance and feedback. Section 4 presents the proposed framework, detailing its functional areas, alignment with the Plan–Execute–Improve cycle, associated controls (governance, approvals, traceability), and implementation artifacts (metrics, quality gates, dashboards). Section 5 reports the case study from the smart-home sector, including study design, KPIs, and comparative results between conventional and generative AI-assisted QA workflows. Section 6 concludes with key findings, implications for policy and compliance, and directions for future research and scaling.

2. State-of-the-Art: Generative AI Tools and Their Applications for Manufacturing Quality Assurance

The use of generative AI in manufacturing quality management is rapidly evolving, supported by advances in natural language processing, computer vision, and multimodal modeling. Generative AI tools are introducing new capabilities across the QA lifecycle – from automated documentation and real-time decision support to compliance reporting and audit preparation. This section reviews the state of the art by: 1) positioning generative AI within QA concepts and industrial use cases; 2) summarizing the principal quality metrics, indices, models, and standards that guide measurement and compliance; and 3) presenting a taxonomy of generative AI-enabled QA tools, their core capabilities, and typical implementation constraints.

2.1. Generative AI in Quality Assurance Contexts

Generative AI refers to AI systems that learn patterns in data and produce new, context-relevant outputs such as text, images, or code. NIST defines generative AI as “the class of AI models that emulate the structure and characteristics of input data in order to generate derived synthetic content” [12]. In manufacturing quality assurance, generative AI typically appears as software platforms that integrate pretrained LLMs and related generative models into quality workflows for planning, process control, documentation, and auditing. Such systems can interpret natural-language requests, draft or revise QA documentation to align with standards, support inspectors with context-aware guidance, and help users navigate complex quality records. Compared with earlier rule-based expert systems, generative AI is more adaptive: rather than following static checklists, a QA assistant can respond to an engineer’s free-form query about a deviation and generate a tailored explanation or a structured action plan aligned with relevant requirements.

The scope of generative AI in QA should be specified carefully. It does not replace established automation such as Statistical Process Control (SPC), conventional machine vision, or Manufacturing Execution Systems (MES) logic; instead, it complements Quality 4.0 environments by working with unstructured information, supporting cross-document reasoning, and improving human decision support. Typical outputs include structured text (e.g., reports, control-plan updates, audit summaries), synthetic examples (e.g., defect illustrations or stress-test scenarios), and code snippets (e.g., scripts to query quality databases). Because generative models learn from large corpora and historical records, they must be constrained and adapted to quality-domain requirements to remain accurate and auditable. Accordingly, NIST guidance emphasizes that generative AI used in high-stakes settings should satisfy trustworthy AI principles such as validity, transparency, and accountability [12,13]. In QA contexts, “generative” therefore implies not only producing content, but producing content that remains verifiable, relevant, and traceable within regulated quality systems.

2.2. Key Quality Metrics, Indices, Models and Standards in Manufacturing Enterprises

This subsection synthesizes the key quality metrics, evaluation indices, reference models, and standards that form the foundation for assessing how generative AI can be integrated into QM.

Quality Metrics and Indices

Manufacturing enterprises employ quantitative metrics and indices to monitor quality performance at organizational and operational levels. These measurements are central to diagnosing inefficiencies, meeting regulatory and customer requirements, and supporting continuous improvement. Commonly used product- and process-level indicators include:

- First Pass Yield (FPY) – measures the percentage of products manufactured correctly without rework [14].
- Defects per Unit (DPU), Defects Per Million Opportunities (DPMO), and Parts Per Million (PPM) – normalize defects across production scales [13,16,17].
- Process Capability Indices (Cp/Cpk) – reflect how well a process meets specification limits [18].
- Overall Equipment Effectiveness (OEE) – evaluates availability, performance, and quality effectiveness of equipment [19].
- Cost of Poor Quality (CoPQ) – aggregates costs from scrap, rework, and warranty claims [20].

These metrics are increasingly monitored through dashboards and may be combined into composite indices supporting operational decision-making. For instance, Lean Quality Control approaches track FPY, OEE, and CoPQ to evaluate operational stability and waste reduction [21]. With IoT data capture and predictive maintenance, early signals of deterioration (e.g., declining Cp/Cpk or OEE) can be detected earlier and acted upon before nonconformities propagate.

Quality Models and Methodologies

Metrics are operationalized through structured methodologies aimed at continuous improvement. The PDCA (Plan–Do–Check–Act) cycle underlies many quality management systems and informs phases of planning, execution, monitoring, and control [22]. Other influential methodologies include:

- Six Sigma – uses the DMAIC (Define–Measure–Analyze–Improve–Control) framework to reduce defects. AI and Machine Learning (ML) methods increasingly support analysis and diagnosis [23].
- Lean and Lean Six Sigma (LSS) – applies tools such as 5S, value stream mapping, and error-proofing (poka-yoke) to reduce waste and variability [24].
- Statistical Process Control (SPC) – maintains stability via control charts; ML-enhanced SPC supports dynamic anomaly detection [25].
- Failure Mode and Effects Analysis (FMEA) – a core tool for proactive risk management; generative AI can support failure identification and action prioritization [26].
- Total Quality Management (TQM) and Root Cause Analysis – are long-established approaches increasingly augmented by digital analytics and data-driven systems [27].

These methodologies define the logic and structure within which generative AI must operate to remain aligned with established quality objectives and to produce outputs that are acceptable for internal governance and external audits.

Quality Standards and AI Integration

Quality standards institutionalize best practices and provide compliance frameworks. In manufacturing, ISO 9001 remains foundational, specifying Quality Management System (QMS) requirements based on the process approach and PDCA logic [28]. Sector-specific extensions add rigor for high-consequence environments; for example, IATF 16949 supplements ISO 9001 with stringent quality planning and defect prevention requirements and emphasizes core tools (e.g., Advanced Product Quality Planning – APQP, Part Approval Process – PPAP, Measurement System Analysis – MSA, SPC) [29,30]. Laboratory-centric standards such as ISO/IEC 17025 impose requirements on technical competence, measurement traceability, and validated methods [31].

Alongside classical quality standards, AI-specific governance standards are increasingly relevant to QA and auditing. ISO/IEC 42001 establishes requirements for an AI management system, including governance structures and risk controls for AI-enabled operations [32]. Supporting

standards and auditing guidance (e.g., measurement management and auditing guidelines) further help operationalize measurement integrity, evidence quality, and audit consistency.

In practice, organizations often align these requirements through integrated management systems that harmonize structures and controls across quality, safety, security, and AI governance [28–33]. These metrics, models, and standards provide the analytical and procedural baseline for evaluating both the benefits and risks of deploying generative AI in quality-intensive manufacturing contexts.

2.3. Taxonomy of Generative AI Tools for Quality Assurance in Manufacturing Companies

Generative AI tools for manufacturing QA can be categorized by function and user role, reflecting their application across the quality lifecycle – from supplier qualification and production control to compliance auditing and governance. Table 1 summarizes the main tool categories, their primary functions, typical users, and representative applications.

Document generation assistants use generative AI to draft and maintain structured QA documentation such as control plans, Process Failure Mode and Effects Analysis (PFMEAs), work instructions/Standard Operating Procedures (SOPs), inspection plans, and standardized reports. Given inputs such as bills of materials (BOMs), process flows, Critical-to-Quality characteristics (CTQs), or prior document versions, these tools can generate consistent drafts aligned with internal templates and external requirements, reducing cycle time in New Product Introduction (NPI) and change-control scenarios [4].

Corrective and Preventive Action (CAPA) narrative generators transform structured inputs (root causes, containment/correction/prevention actions, verification results) into traceable CAPA narratives suitable for internal reviews and external audits. Their value is strongest when they preserve linkage between nonconformance evidence and closure criteria, and when outputs follow predefined schemas that support audit defensibility [34].

Supplier QA portals with generative AI streamline supplier communication and document exchange by supporting intelligent form completion, automated feedback on submissions (e.g., PPAP evidence completeness), and supplier-specific scorecards. When integrated with supplier history and specifications, these tools can reduce back-and-forth iterations and improve “first-time-right” submission quality [11,35].

Interactive inspection agents operate near the shop-floor by providing context-aware guidance during inspection routines. Using chat or voice interfaces, they can clarify procedures, retrieve relevant work instructions, and help operators interpret borderline results. Their effectiveness depends on safe integration with validated procedures and clear guardrails for advisory outputs [4].

Anomaly explanation engines generate interpretable hypotheses for anomalies detected in SPC trends, sensor signals, or inspection outcomes. Rather than replacing detection systems, they add interpretability and decision support by converting heterogeneous evidence (process context, historical Non-Conformance Reports (NCRs), equipment conditions) into structured explanations and candidate corrective actions [36].

Compliance and audit review tools support audit readiness by checking record completeness, identifying documentation gaps, and mapping evidence to audit clauses. These tools can generate audit checklists, clause-to-evidence summaries, and pre-audit packages, reducing manual preparation burden while increasing consistency of evidence presentation [37].

Governance dashboards provide oversight across AI-assisted QA activities by consolidating recommendations, approvals, action status, and traceability signals. By summarizing AI outputs and surfacing exceptions (e.g., unresolved actions, missing approvals, weak evidence chains), they support management review and ongoing compliance monitoring [38].

Across these categories, generative AI acts primarily as a linguistic and reasoning layer that translates complex QA data into auditable and actionable content. However, implementation constraints are non-trivial. High-stakes QA deployments require: 1) controlled data access and role-based permissions, 2) verifiable grounding to authoritative sources (e.g., retrieval with citation

enforcement), 3) version control and change traceability, and 4) human-in-the-loop review for decisions with compliance or safety implications [12,13]. Consequently, tool selection and integration should align with process criticality, data maturity, and governance capability. Documentation and audit-support tools often yield early benefits with modest integration demands, whereas real-time inspection and anomaly-explanation tools require stronger instrumentation, validation discipline, and organizational change management.

Table 1 consolidates these categories and maps each to its primary function, typical users, and representative applications.

Table 1. Summary of generative AI tools in manufacturing quality assurance.

Tool Category	Primary Function	Primary Users	Representative Applications
Document Generation Assistants	Create and update QA/QMS documentation (control plans, SOPs, WIs)	Quality engineers; QA managers	Draft control plans or work instructions using prior versions and product specifications
CAPA Narrative Generators	Generate structured narratives for NCRs and CAPA records	Compliance officers; auditors	Produce RCA/CAPA narratives linked to inspection evidence and traceable quality records
Generative AI-enabled Supplier QA Portals	Standardize supplier submissions and guide required documentation	Suppliers; supplier quality/procurement teams	Validate PPAP completeness and flag inconsistencies with specifications
Interactive Inspection Agents	Provide real-time guidance during inspections and quality checks	Line operators; floor supervisors	Voice/chat assistants that guide steps, answer procedure questions, and log results
Anomaly Explanation Engines	Explain anomalies and propose corrective actions	Process engineers; QA analysts	Summarize abnormal trends from sensor/SPC data and suggest likely causes
Compliance/Audit Review Tools	Identify gaps and compile audit-ready compliance evidence	QA leads; internal auditors	Review QMS records against standards and generate checklists or regulator-ready summaries
Governance Dashboards	Monitor AI outputs, approvals, and traceability for oversight	QA directors; risk managers	Aggregate AI recommendations, human approvals, and open issues for management review and audits

Each tool category is applied in a different operating context and supports distinct QA functions. Documentation-oriented tools (e.g., document assistants and CAPA narrative generators) are commonly used in offline workflows to accelerate drafting and reporting while preserving required formats. Operational tools (e.g., interactive inspection agents and anomaly explanation engines) deliver real-time guidance on the shop floor or in engineering control rooms, where rapid interpretation of deviations and immediate response are critical. Oversight tools (e.g.,

compliance/audit review tools and governance dashboards) operate at a supervisory layer, connecting to enterprise quality systems to consolidate evidence, surface gaps, and maintain auditable decision trails. In practice, many generative AI capabilities are embedded into MES, Product Lifecycle Management (PLM), and QMS platforms to augment human-led quality processes with contextual assistance; adoption is especially visible in documentation-heavy sectors where CAPA drafting and evidence compilation directly support audit readiness (e.g., automotive and aerospace).

This role-anchored taxonomy highlights four functional groupings:

- Authoring and compliance documentation tools (document assistants; CAPA narrators) support structured QMS records and reporting. Because these outputs can affect compliance, human-in-the-loop review remains essential to confirm accuracy, traceability, and standards alignment.
- Supplier-facing tools (generative AI-enabled supplier QA portals) improve submission consistency and first-time-right rates by guiding suppliers through required evidence and checks. Their deployment requires strong validation and data-protection controls due to proprietary supplier inputs.
- Operational guidance tools (interactive inspection agents; anomaly explainers) assist frontline personnel with procedure retrieval, deviation interpretation, and context-aware troubleshooting. Their effectiveness depends on reliable integration, low latency, and user trust calibrated by clear guardrails.
- Oversight and audit tools (compliance/audit review tools; governance dashboards) support system-level governance by verifying the completeness of records, monitoring approvals, and tracking unresolved actions in line with ISO 9001 [28] and IATF 16949 [29].

A key operational distinction follows from this typology: some tools act primarily as assistive interfaces for documentation and audit preparation, whereas others interact more directly with control and verification loops. Integration strategies should reflect this difference. Assistive tools benefit from structured workflows such as redlining, template enforcement, and version control to make outputs reviewable and reproducible. Tools that influence control decisions require stronger safeguards, including embedded explainability, escalation logic, and clearly defined accountability for AI-supported recommendations, especially where MSA-related constraints apply.

Implementation success depends on matching tool capabilities to process criticality, data readiness, and organizational maturity. Document assistants and audit review tools often deliver early value with limited integration effort, making them suitable entry points. In contrast, real-time inspection guidance and anomaly explanation typically require richer instrumentation, domain adaptation, and change management to ensure stable operation and user adoption. Governance dashboards play a cross-cutting role by aggregating outputs across systems and enabling end-to-end visibility over AI-supported QA activities.

These tool categories form a layered ecosystem that extends from document generation to audit traceability and supports hybrid human–AI quality architectures grounded in verifiability, accountability, and regulatory compliance.

3. Related Work

Rapid progress in generative AI has accelerated both academic research and industrial experimentation on its use in manufacturing quality assurance. This section reviews the most relevant work along two complementary streams: 1) applications of generative AI tools that augment QA activities across the production lifecycle, and 2) conceptual and operational frameworks that govern the responsible integration of generative AI into quality management systems. These studies motivate our research and clarify unresolved challenges that the proposed framework is designed to address.

3.1. Applications of Generative AI Tools for Quality Assurance in Manufacturing Companies

Over the last decade, shop-floor and laboratory quality assessment has moved from paper-based records and isolated SPC tools toward integrated digital environments that connect SPC, MSA, eQMS, MES, and PLM. This shift has enabled systematic data capture, risk-based sampling, and closed-loop dashboards. The most recent transition is the introduction of generative and multimodal AI capabilities that can draft controlled documentation, explain deviations, link evidence across systems, and compile audit-ready narratives – provided outputs remain source-grounded and are formally approved within the QMS before release [39]. In parallel, predictive-quality and digital-twin research provides methods and performance baselines for evaluating impact on operational outcomes such as FPY, PPM/DPMO, Cp/Cpk, CAPA timeliness, and audit nonconformities [40,41].

Below, we group prominent classes of generative AI-enabled assistants according to the assessment needs they address, highlighting typical benefits and the assurance controls required for safe deployment.

- QMS/MES/PLM-embedded copilots (controlled drafting and evidence linkage): – Copilots embedded in eQMS, MES, and PLM can generate initial drafts of PFMEA sections, control-plan updates, PPAP summaries, and work instructions by grounding on CTQs, specifications, SPC/MSA results, and historical NCR/CAPA records. Reported benefits include shorter revision cycles and more consistent documentation, while key controls include enforced citations to approved sources, versioned audit trails, and human sign-off prior to release [42].
- Vision-language assistants for automated optical inspection (AOI) and in-process diagnostics: – Multimodal models can interpret AOI imagery (or line-camera outputs) together with SPC signals to produce operator-facing explanations (e.g., defect pattern + chart shift suggesting paste-volume instability). Evidence from industrial anomaly-detection research – including recent surveys and diffusion-based methods – indicates substantial gains in image-based detection performance; in QA settings, these models are typically positioned as advisory and require MSA-aligned validation and periodic requalification [43,44].
- Retrieval-augmented generation (RAG) systems: – answer free-text questions (e.g., reaction plans after a rule violation) by retrieving approved work instructions, control plans, and customer standards, then generating responses with explicit citations. Recent smart-manufacturing studies show that hybrid RAG designs – combining metadata/knowledge-graph structure with vector retrieval – improve precision and traceability, making them suitable for controlled QA knowledge support [45].
- Supplier quality and incoming-inspection copilots: – Assistants can summarize supplier defect histories (e.g., NCR patterns and PPM trends), flag documentation gaps, and propose sampling adjustments based on risk indicators. While these functions combine conventional analytics with text generation, governance typically requires that dispositions remain under Supplier Quality Engineer responsibility, supported by access control and traceable rationales [42].
- CAPA and audit-evidence assemblers: – Generative tools can draft 8D narratives, map containment/correction/prevention actions to PFMEA causes, and compile evidence packs (logs, test results, photos) using templates and schema validation with explicit role approvals. Related conformance-checking and process-mining research demonstrates how event logs can be transformed into objective audit evidence and KPIs for audit readiness, providing structures that generative systems can leverage to improve documentation completeness and speed [46].
- Metrology and laboratory reporting aides (ISO/IEC 17025 contexts): – In testing and calibration laboratories, assistants can support drafting of method descriptions, uncertainty narratives, and result interpretation summaries. However, these outputs must remain consistent with validated methods, documented uncertainty budgets, impartiality requirements, and authorized sign-off practices expected in accredited environments [47].
- Simulation and digital-twin scenario generators: – By coupling process models with generative AI, organizations can explore “what-if” scenarios for inspection planning (e.g., changing sampling regimes) and generate stress-test cases before releasing process changes. Recent work

also shows LLM-enabled digital-twin approaches that learn temporal features from production data, supporting rapid hypothesis testing, confirmation runs, and evidence-backed parameter updates [41].

Across these categories, recurring assurance controls include: human-in-the-loop approvals for controlled outputs; strict source grounding with citations; segregation of duties in audit workflows; role-based access control and data minimization; monitoring for drift with periodic revalidation; and documented change control for prompts, templates, and retrieval indexes. Evaluation should prioritize operational and compliance outcomes – FPY, PPM/DPMO, Cp/Cpk, CAPA lead time, and audit findings – to verify that AI strengthens assessment quality rather than introducing new failure modes. Independent evaluations further highlight current limitations of vision–language and multimodal LLMs on industrial tasks, reinforcing the prevailing practice of deploying them as decision support rather than as autonomous control in regulated QA [48].

Table 2 provides a comparative overview of widely adopted categories of generative AI tools in manufacturing QA, summarizing their typical tasks, data dependencies, system integrations, and assurance mechanisms to clarify how each tool type fits within the broader quality ecosystem.

Table 2. Comparison of widely used generative AI tool categories for manufacturing quality assurance.

Category	Typical QA tasks	Primary data	Integrations	Example outputs	Key assurance controls	Where it fits
QMS/MES/PLM co-pilot [42]	Draft PFMEA, control plans, PPAP packages, work instructions, summarize NCRs	CTQs, specifications, NCR/CAPA history	eQMS, MES, PLM	Drafted sections with citations, change logs	Human approval, version control, source-locked RAG	Supplier quality, in-process QA
Vision–language AOI aide [43,44]	Explain defects, recommend inspection checks	AOI/line images, SPC signals	AOI, MES	Defect rationales, operator checklists	MSA-aligned validation, advisory-only use	In-process QA
RAG QA assistant [45]	Line-side Q&A with citations, retrieve reaction plans PPAP	Approved work instructions, control plans, standards	eQMS, DMS	Grounded answers with links	Access control, citation enforcement	In-process QA, audit support
Supplier quality co-pilot [42]	completeness checks, risk-based sampling proposals	Supplier performance history, specifications	ERP, SQM, eQMS	Risk-stratified sampling plans, scorecards	SQE approval, auditable rationale trail	Supplier quality, incoming inspection
CAPA/audit assembler [46]	Draft 8D/CAPA narratives, compile evidence packs	NCRs, test results, event logs	eQMS, LIMS	Structured reports, clause-to-evidence matrices	Schema validation, role-based sign-off	Cross-functional (QA, production, audit)
Metrology aide [47]	Draft method descriptions and	Lab methods, measurement data,	LIMS, QMS	ISO/IEC 17025-aligned narratives	Impartiality controls, authorized signatories	Testing, release decision support

	uncertainty narratives	uncertainty budgets				
	What-if studies,	CTQs/specs,		CTQ/yield predictions		Pre-production,
Simulation and digital-twin aide [41]	DOE support, Cp/Cpk and yield forecasts	SPC, NCR/CAPA, sensor/process data	MES, SPC/eQMS, PLM/CAD	, sensitivity drivers, operating limits	V&V, input lineage, human approval	qualification, in-process optimization

Note: The abbreviations not previously mentioned in the text can be found in the Abbreviations section.

Generative AI is rapidly transitioning from a novel concept to an operational cornerstone in manufacturing quality assurance. As shown in Table 2, these tools span a wide range of use cases – from drafting PFMEAs and control plans to supporting line-side inspections and assembling audit evidence. Each tool class integrates specific data sources and systems (e.g., eQMS, MES, PLM, LIMS) and is governed by targeted assurance controls such as human-in-the-loop validation, citation enforcement, or versioning protocols. This alignment ensures that generative AI enhances, rather than compromises, quality integrity across supplier, in-process, and cross-functional domains.

Reported implementations in documentation-heavy and highly regulated sectors suggest practical benefits, particularly in documentation throughput, evidence structuring, and decision support during deviations. In most deployments, generative AI acts as a “copilot” that augments expert judgment rather than replacing it, with controlled outputs released only after review and approval. Accordingly, successful adoption depends on explicit governance procedures and alignment with established QMS requirements (e.g., ISO 9001 [28] and, where applicable, IATF 16949 [29]). As tool maturity increases, integration is expected to expand toward richer interfaces (e.g., conversational line-side support) and more automated evidence packaging, while remaining anchored in transparency, traceability, and accountability.

3.2. Frameworks for Application of Generative AI in Manufacturing QA

In parallel with tool development, recent studies have proposed frameworks that structure how generative AI can be embedded in manufacturing QA, including system architecture, functional scope, and safeguards (Table 3). This subsection reviews representative approaches, emphasizing how they address integration with quality processes, evidence traceability, and risk controls, and it positions the operational framework proposed in Section 4.

Recent work illustrates several directions. Rydzi et al. [49] developed a predictive inspection framework for automotive production that uses machine learning to triage end-of-line inspections based on defect risk. Although generative components were not central, the authors note the potential role of language models in organizing recurring warranty failure patterns, highlighting the value of integrating AI into closed-loop quality improvement. Nguyen et al. [50] proposed XedgeAI, a human-centered inspection framework that combines explainable AI with vision–language models for edge-deployed defect detection and incorporates expert feedback to improve interpretability in mobile or distributed inspection settings. Lin et al. [51] introduced DDD-GenDT, a dynamic data-driven digital twin architecture that uses LLM ensembles to support real-time adaptation (e.g., tool-wear prediction), illustrating how generative reasoning can complement monitoring under time and data constraints.

A related research stream focuses on using generative models to improve QA data availability and automation. Shafiee [52] reviewed the use of generative models (e.g., GANs and VAEs) for synthetic defect image generation to strengthen visual inspection pipelines. Thomas [53] and Alsaif et al. [54] reported LLM-based approaches for accelerating FMEA drafting and supporting fault diagnosis. Wang et al. [55] proposed a multimodal LLM framework to support shop-floor decision-making via real-time question answering for inspection and maintenance tasks.

Despite these advances, many published frameworks remain scoped to specific domains or functions and provide limited guidance on end-to-end integration within a QMS. For example, Rydzi et al. [49] focus on end-of-line inspection triage, XedgeAI [50] emphasizes edge-based vision inspection, and DDD-GenDT [51] is oriented toward adaptive digital twins. Cross-cutting requirements – such as traceable evidence packaging, role-based approvals, change control, and audit readiness – are often treated implicitly or left to implementation. Shadid et al. [56] similarly observe that many Quality 4.0 frameworks are conceptual, resource-intensive, and insufficiently validated in real operational deployments.

More recent approaches move toward tighter integration of foundation models with controlled enterprise knowledge and governance. Álvaro and González Barreda [57] presented a retrieval-augmented generation system for QA that answers operator questions using curated document repositories. Wan et al. [58] extended this direction by incorporating knowledge graphs to improve factual consistency and traceability of generative outputs. Complementary work by Sun et al. [59] and Mata et al. [60] positions LLMs to support adaptive monitoring and evidence-backed adjustments as process states evolve, including within digital twin ecosystems. Across these contributions, a common trend emerges: scalable manufacturing QA frameworks require not only capable foundation models, but also controlled data access, auditable outputs, and explicit QMS-aligned safeguards.

Table 3. Overview of recent frameworks for applying generative AI in manufacturing quality assurance.

Reference	GAI tool/ component	Target QA feature	Methodological approach
Rydzi et al. (2024) [49]	Predictive inspection triage (ML, GenAI proposed)	End-of-line defect prediction	ML-based triage with planned GenAI extension
Nguyen et al. (2024) [50]	XedgeAI (XAI + LVLMS)	Edge-based visual inspection	Modular XAI system with data augmentation
Lin et al. (2025) [51]	DDD-GenDT (LLM-augmented DT)	Real-time process monitoring	LLM ensemble in adaptive digital twin
Shafiee (2025) [52]	GAN/VAE for defect synthesis	Visual model training augmentation	Synthetic data generation
Thomas (2023) [53]	ChatGPT for FMEA drafting	Failure mode analysis	Prompt-based failure scenario generation
Alsaif et al. (2024) [54]	Multimodal LLM	Fault detection and diagnosis	Signal-text LLM fusion for anomaly diagnosis
Wang et al. (2025) [55]	LLM-based Q&A assistant	Real-time decision support	Shop-floor assistant with multimodal input
Álvaro & González (2025) [57]	RAG assistant for QA	Line-side Q&A over controlled knowledge	Controlled retrieval with versioned sources
Wan et al. (2025) [58]	Hybrid KG+RAG	Factory Q&A and guidance	Ontology-constrained RAG for QA
Sun et al. (2024) [59], Mata et al. (2025) [60]	LLM, LLM-enhanced digital twins	Dynamic inspection and confirmation	LLM integration into live simulation pipelines

The reviewed frameworks demonstrate multiple pathways for applying generative AI in manufacturing QA, typically addressing isolated functions such as line-side guidance, visual inspection support, or model-assisted monitoring. Together, they show growing technical maturity in components such as RAG, multimodal models, and LLM-augmented digital twins, but they also expose recurring gaps in governance, traceability, and standards alignment that limit audit defensibility and enterprise-scale adoption.

RAG-based assistants and Knowledge Graph (KG)+RAG approaches improve access to controlled knowledge and can accelerate decision support, yet their reliability depends on retrieval from approved, version-controlled sources, strict access governance, and mandatory human validation. Digital twin-oriented frameworks further enable adaptive monitoring and scenario-

driven planning, but they intensify traceability demands because model logic and operating conditions evolve over time, requiring lifecycle logging, change control, and periodic re-validation. Similarly, explainable inspection systems improve transparency at the point of use, but they still require controlled storage, versioning, and explicit linkage to source data and approvals to meet audit expectations.

These limitations motivate our multi-level, multi-branch framework. A single tool-centric architecture is insufficient because QA evidence and decision rights are distributed across: 1) hierarchical assessment levels – product, process, and operation – and 2) distinct stakeholder branches – manufacturers, auditors, and oversight/control entities. Our framework therefore structures generative AI support across these levels and branches within a common five-phase QA cycle, so that generated outputs remain: a) grounded in controlled sources, b) governed through role-based approvals and segregation of duties, and c) traceable end-to-end from operational signals to controlled documentation and audit conclusions. This design closes the gap between technically capable prototypes and production-ready, standards-aligned QA infrastructure.

4. Framework for Generative AI-Supported Quality Assurance in Manufacturing Enterprises

This section introduces a structured framework for embedding generative AI tools – LLM-based assistants, semi-autonomous agents, and analytics components – into industrial QA workflows. The aim is to support a scalable and auditable QA ecosystem by integrating AI capabilities into both manufacturing and assurance activities within a unified, hierarchical architecture.

The framework addresses growing requirements for quality, compliance, and operational efficiency by coordinating three participant groups: manufacturers (left branch), auditors (right branch), and system-level QA oversight/control entities (top branch). Generative AI supports each group with role-specific functions while maintaining clear accountability through governed inputs, approvals, and traceable outputs.

As shown in Figure 1, the framework is deployed across three nested assessment levels – product-level, process-level, and operation-level assurance. The QA algorithms is organized into a common five-phase cycle (planning, pre-run preparation, in-process monitoring, post-run correction, and review). Within this structure, generative AI strengthens real-time visibility, documentation consistency, and response speed, while preserving human oversight and end-to-end traceability.

The three levels define the scope of assessment and the granularity of evidence required. As the framework progresses from product-level to process- and operation-level assurance, the emphasis shifts from evaluating individual units or batches to explaining process variation and to assessing overall system performance. In parallel, generative AI extends from supporting inspection and defect handling to enabling evidence synthesis, documentation standardization, and traceable, auditable decisions across stakeholders.

Level 1. Product-Level Assessment

At this level, QA focuses on discrete units or individual batches. Manufacturers use AI to define inspection parameters, simulate failure modes, and generate tailored quality checklists from historical data and design specs. During execution, AI agents monitor anomalies in sensor and vision data, adjust sampling dynamically, and support frontline teams with chatbot-guided decisions. Post-run, generative AI clusters defects, drafts CAPA reports, and produces structured feedback. Auditors benefit from real-time access to conformance logs and anomaly reports, allowing timely verification of product integrity and decision traceability.

Level 2. Process-Level QA

Here, the framework shifts focus to production lines and process stability. Generative AI analyzes past performance and suggests control limits, identifies deviations in live SPC charts, and validates machine configurations against standards. Throughout the production run, audit bots monitor procedural compliance, highlight process drifts, and document deviations. After completion, AI-driven analysis correlates defects with process inputs, enabling root cause discovery and CAPA

validation. Results feed into broader process improvement cycles and influence future inspection strategies.

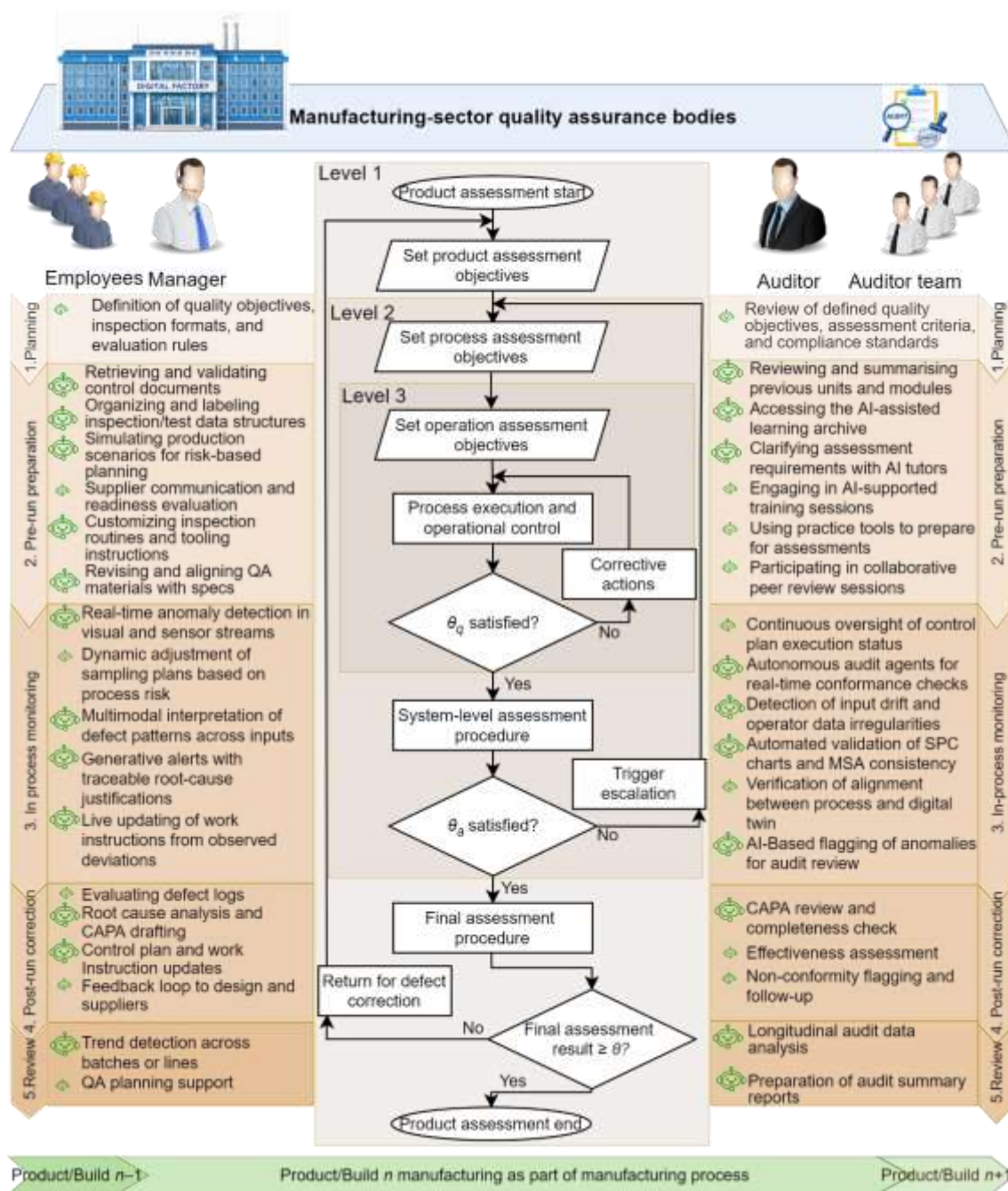


Figure 1. Framework for generative AI application in Manufacturing QA. Note: The size of the green icons representing generative AI tools (left and right branches) indicates the relative extent of generative AI applicability to each quality assessment activity.

Level 3. Operation-Level Assurance

At the operational level, the framework supports holistic oversight of an entire manufacturing facility. Generative AI synthesizes data from all production lines, assesses global compliance metrics, and drafts plant-wide quality policies informed by benchmarking and standards. Central AI systems coordinate audit readiness, simulate risks, and propose systemic improvements. During operation, orchestration agents maintain a real-time digital twin, enabling predictive risk detection and inter-process insights. Strategic reviews aggregate performance data over time to guide investment, training, and operational excellence initiatives.

As shown in Figure 1, the proposed framework is represented as a central QA decision workflow positioned between the manufacturer-facing activities on the left (factory branch) and the auditor-facing activities on the right (audit branch). It operates across three nested assessment levels – product-level, process-level, and operation-level assurance – while all participants follow the same five-phase cycle: planning, pre-run preparation, in-process monitoring, post-run correction, and review. Within this shared structure, generative AI supports faster sense-making and more consistent documentation, but decisions remain auditable through human oversight and end-to-end traceability.

The middle workflow begins once product assessment objectives are defined and progressively expands the scope of evaluation from the product to the process and then to the operation level, reflecting the increasing depth of evidence required. During execution and monitoring, manufacturing performance is checked against the manufacturing-quality threshold, θ_q , which is defined using key quality metrics (FPY \geq 95%, DPMO \leq 500 \approx 500 PPM, and Cp/Cpk \geq 1.33). If θ_q is not met, corrective actions are initiated to restore conformance before proceeding. When manufacturing quality is satisfactory, the workflow advances to a system-level assessment, where audit readiness is evaluated against the audit-assurance threshold, θ_a , requiring 0 major nonconformities, at most 2 minor nonconformities, CAPA on-time closure \geq 90%, and fully verified traceability and data integrity (i.e., complete, tamper-resistant records). If θ_a is not satisfied, escalation is triggered, meaning the case is formally elevated for higher-level review with documented follow-up and accountability. If audit assurance is satisfied, the process proceeds to the final assessment stage, where the overall acceptance score is compared to θ , an overall passing threshold derived from the applicable certification or assessment scheme (in practice, often around 60–70% of total points). If the final score does not reach θ , the workflow returns for defect correction and re-entry into the improvement loop; otherwise, the assessment concludes. The thresholds θ_q , θ_a , and θ are product-specific and are formally set during the pre-run preparation phase. They are derived from design specifications, applicable regulatory requirements, and historical performance data to ensure that quality and audit evaluations are calibrated to the manufacturing context.

This framework also outlines how generative AI enhances QA processes in industrial environments. It mirrors traditional QA workflow steps and maps the application of generative AI tools to the respective responsibilities of manufacturers and auditors. Each step fosters traceability, responsiveness, and continuous improvement.

Step 1: Planning

In the planning phase, manufacturers define quality objectives, inspection formats, and evaluation rules with support from generative AI-driven historical analysis, defect pattern prediction, and standards-based planning tools. This ensures alignment of control expectations with prior performance and design intent. Simultaneously, auditors review and validate these definitions using AI to cross-reference industry benchmarks, simulate failure scenarios, and assess compliance frameworks. AI helps both parties converge on clear, measurable goals that form the baseline for subsequent QA actions.

Step 2: Pre-Run Preparation

Manufacturers use generative AI to retrieve and validate control documents, organize inspection data structures, simulate production runs for risk-based prioritization, and assess supplier readiness through automated credential reviews and communication. AI-driven planning agents also generate or refine inspection routines and tooling specifications. On the auditor side, generative AI supports the review of prior audit findings and nonconformance records, highlights common deficiencies, and ensures that control documentation is up-to-date. Interactive AI checklists and training modules help auditors prepare efficiently, while benchmarking tools ensure that audit scope aligns with regulatory and corporate expectations.

Step 3: In-Process Monitoring

During manufacturing execution, generative AI enables real-time anomaly detection by interpreting visual and sensor data streams, dynamically adjusting sampling rates based on process risk, and issuing traceable alerts tied to specific failure modes. It can update work instructions in

response to detected deviations and provide decision support via line-side AI chatbots. For auditors, autonomous agents continuously monitor control plan adherence, validate operator inputs, and verify that SPC and MSA procedures remain within acceptable thresholds. AI further detects drift, ensures consistency with digital twin models, and flags anomalies for interim review.

Step 4: Post-Run Corrections

After production, manufacturers analyze defect logs, perform root cause analysis, and draft CAPA reports. Generative AI helps generate NCR documentation, update control plans, and feed corrective insights back into upstream design or process steps. It also compiles operator-level feedback for broader learning. Auditors, supported by generative tools, evaluate CAPA completeness, verify corrective effectiveness, and request clarifications or supporting evidence as needed. AI-generated summaries and traceability logs ensure completeness and transparency, while systemic recommendations are flagged to drive long-term quality improvements.

Step 5: Review

In the final phase, manufacturers use generative AI analytics to detect cross-batch trends, identify chronic issues, and guide strategic QA planning. These insights contribute to refining inspection strategies and mitigating recurring risks. Auditors engage in longitudinal analysis of accumulated audit data, using AI to produce comprehensive summary reports that track improvements or degradations in process maturity over time. The outputs inform both compliance assurance and future audit planning cycles.

This generative AI-enabled industrial QA framework aligns technological innovation with quality assurance rigor, supporting all actors in the manufacturing ecosystem – from production engineers and QA professionals to internal and external auditors. Designed for flexibility and adaptability across diverse manufacturing contexts, it enables a balanced integration of automation and expert oversight. The framework is structured to reflect five core QA phases – planning, pre-run preparation, in-process monitoring, post-run correction, and review – applied across three hierarchical levels of quality control: product-level, process-level, and operation-level assessments.

For manufacturers, the framework ensures timely QA interventions, minimizes waste through early defect prediction, and increases efficiency through automation. For auditors, AI enhances transparency, allows real-time access to structured compliance evidence, and supports consistency across evaluations. By enabling proactive, data-informed action, this model fosters a more resilient, responsive, and explainable QA system.

Moreover, by embedding AI into a clearly phased assessment cycle, the framework promotes industrial QA principles such as consistency, adaptability, continuous improvement, and traceability. It supports personalized quality goals at the product level, dynamic oversight at the process level, and strategic coordination at the operational level. The integration of generative AI strengthens systemic integrity while empowering human actors with the tools to achieve higher precision, compliance, and productivity.

Additionally, generative AI enhances industrial audit processes by supporting the validation of inspection protocols, fairness in sampling procedures, and alignment with established quality control and compliance frameworks. Natural language processing-drive tools can evaluate whether inspection instructions are overly narrow or ambiguous, detect duplications or inconsistencies in control plan documentation, and flag potential procedural gaps that may compromise traceability or integrity. AI-assisted review of operator comments, maintenance logs, and supplier assessments also informs continuous improvement cycles and supports evidence-based QA auditing. Through these capabilities, internal and external audit bodies gain more structured oversight, faster reporting, and greater consistency in evaluating compliance with regulatory and industry-specific standards.

However, implementing generative AI in manufacturing QA also brings critical challenges that require diligent oversight. Chief among them is the potential for bias in automated quality assessments – such as skewed judgment due to unbalanced training datasets or overfitting to past defects. Such risks may lead to either overreporting issues or overlooking emerging deviations. To address this, regular audits of AI tool performance are essential, including calibration reviews,

transparency in training data origins, and human-in-the-loop validation of non-conformity assessments. Over-reliance on automation must be avoided: engineers and auditors must retain decision-making authority, particularly in interpreting borderline cases and confirming CAPA effectiveness. Operational policies should mandate hybrid oversight models where AI-generated outputs serve as decision support tools, not final arbiters.

This generative AI-driven framework introduces a robust, tiered approach to industrial QA, integrating innovation while safeguarding process credibility and human accountability. By embedding intelligent automation into inspection design, execution, and audit review, the framework enhances visibility, responsiveness, and documentation across the product, process, and operational levels. It accommodates a broad range of generative AI applications, from predictive analytics and automated anomaly detection to report generation and digital twin validation.

To demonstrate how the proposed framework works in practice, Section 5 applies it in a bounded proof-of-concept (PoC) New Product Introduction (NPI) quality-planning task. The PoC operationalizes the framework's assessment logic through explicit objectives and decision gates – manufacturing quality θ_q , audit assurance θ_a , and the overall acceptance threshold θ – and reflects the two role branches by requiring manufacturer-side planning artifacts and auditor-oriented evidence packs. Both the human-only and the generative-AI-assisted conditions follow the same five-phase QA cycle and the same governance requirements (document control, approvals, and traceability), allowing the study to test whether generative AI improves outcomes within the framework rather than replacing established quality controls. In addition, the validation procedure provides a practical template for quality professionals, plant engineers, and audit coordinators to quantify improvements in efficiency, traceability, and audit readiness for a clearly defined workflow.

5. Validation of the Proposed Generative AI-Based Quality Assurance Framework

To assess the practical applicability of the proposed generative AI-enabled industrial QA framework, we conducted a structured PoC case study focused on an NPI quality-planning task. The task was completed under the same 10-business-day constraint in two parallel conditions: 1) a team of three QA experts and 2) an LLM-driven workflow using ChatGPT-5-Thinking. Both conditions were given the same product description, deliverable requirements, and acceptance criteria, enabling a direct comparison of planning efficiency, documentation quality, risk coverage, and audit readiness. As a PoC, the objective is to demonstrate feasibility and comparative value for a representative NPI task rather than to claim generalizability across all products, plants, or regulatory settings.

5.1. Case Study: NPI Quality Planning for Smart Thermostat ST-200

The case concerns quality planning for the ST-200 smart thermostat, which includes an RF mainboard, capacitive touchscreen, temperature sensing module, and an injection-molded enclosure. Within the defined time window, both approaches were required to produce four readiness deliverables: 1) a Process FMEA (PFMEA), 2) a linked Control Plan, 3) PPAP requests/templates for key suppliers, and 4) an evidence package suitable for internal audit review. The study follows the framework's assessment logic by applying two product-specific gates. The manufacturing quality threshold θ_q was defined on CTQs as $FPY \geq 95\%$, $DPMO \leq 500$, and $Cp/Cpk \geq 1.33$. The audit assurance threshold θ_a required zero major nonconformities, no more than two minor nonconformities, CAPA on-time closure $\geq 90\%$, and verified traceability and data integrity. These thresholds were set during pre-run preparation to reflect product risk, complexity, and compliance expectations. If θ_q is not met, corrective actions are initiated within the manufacturing workflow; if θ_a is not met, escalation is triggered through documented higher-level review. Because the PoC focuses on NPI planning, it most directly instantiates the planning and pre-run preparation phases,

while later phases are represented through the specified monitoring logic, evidence structure, and predefined triggers rather than long-duration shop-floor deployment.

5.2. Expert Team Solution

The first solution was developed by a team of three QA professionals with experience in AI-supported manufacturing environments. The team drafted the required planning artifacts, considered low-frequency (rare) failure scenarios, and verified the control logic and reaction plans. Their PFMEA emphasized critical surface-mount technology (SMT), bonding, firmware flashing, and final assembly steps, while the linked control plan focused on Cp/Cpk monitoring for CTQs and fully traceable reaction procedures. All work followed the organization's ISO 9001-aligned QMS procedures, including document control, internal review, and formal approval workflows [28].

Table 4 presents an extract of the expert team's PFMEA, which prioritized high-risk steps in the electronics and final assembly flow, including solder paste printing, component placement, reflow, display bonding, firmware flashing, and functional testing. The control plan derived from this analysis relied on 100% solder paste inspection (SPI), AOI aligned with Acceptability Standard for Electronic Devices (IPC-A-610) criteria, and torque traceability using DC electric torque tools. Overall, the expert documentation was clear and fit for the NPI task, but it contained fewer explicit mechanisms for clause-to-evidence mapping, formalized change logging, and predefined audit-trigger logic that would directly support rapid audit packaging.

Table 4. Extract of PFMEA – expert team.

Step	Failure Mode	Effect	Cause	S/O/D/RPN	Current Control	Action
SMT Print	Solder bridging (RF area)	No connection, high draw	Excess paste, stencil wear	8/4/4/128	SPI, AOI	Reduce aperture, retrain operator
Placement	Tombstoning of 0402s	Intermittent RF issues	Offset, thermal stress	7/3/5/105	AOI, placement program verification; feeder/nozzle checks	Adjust design layout, feeder check
Reflow	Voids under PMIC	Thermal instability	Poor soak profile	9/2/5/90	X-ray, FT	DOE on profile, SPC control
Display Bond	Misalignment > 0.2 mm	Light-bleed	Fixture drift, excess glue	6/4/5/120	Visual check	Add vision inspection, GRR validation
Firmware Flash	Wrong image	Bricked unit	Config control lapse	9/2/6/108	CRC, manual log	Auto flash check, signed build required
Functional Test	Temp offset > ±0.5°C	Inaccurate user comfort	Sensor miscalibration	8/3/4/96	Calibration, FT	Cp/Cpk monitoring on calibration CTQ, drift alarms, supplier sensor CoC/verification

The control plan included methods such as 100% SPI, AOI with IPC-A-610 visual checks, and torque logging using DC tools. While the documentation was clear and task-relevant, limitations were observed in clause-to-evidence mapping, formal change logging, and pre-set audit alarm logic.

5.3. GPT-5-Thinking Solution (Generative AI-Driven Planning)

The second condition applied the same deliverable requirements using ChatGPT-5-Thinking within the proposed framework. The model generated a structured PFMEA and linked control-plan logic, produced PPAP checklist templates for the electronics and plastics suppliers, and assembled an audit-evidence package with embedded traceability references. Table 5 shows an extract of the resulting PFMEA; compared with the expert extract, it emphasizes additional governance-oriented metadata (ownership, residual risk notation, and evidence linking) rather than maintaining a strictly identical set of process steps.

Table 5. Extract of PFMEA – GPT-5-Thinking.

Step	Failure mode	Effect	Cause	S/O/D/RP N	Current control	Recommended action	Owner	rRP N
SMT Print	Solder bridging	Electrical failure	Overpaste, stencil wear	8/4/4/128	100% SPI, AOI confirmation	Optimize aperture design, enforce stencil life/cleaning, retrain operator DOE to optimize	ME-D3	72
Reflow	Voids on PMIC pad	Power instability /shutoff	Inadequate thermal soak	9/2/5/90	Profile verification, X-ray sampling, FT correlation	soak/reflow profile, add SPC limits on key reflow parameters, define reaction plan for void-rate drift Standardize illumination/contrast calibration, add rule-based	PE-D5	54
AOI	False negative (fine-pitch)	Latent field failure	Glare, poor contrast, suboptimal thresholds	8/6/2/96	AOI program control, golden board checks, periodic verification	“high-risk” recheck (e.g., X-ray/ICT sampling), perform MSA/GRR on AOI setup changes	VisEng-D6	48
Firmware Flash	Wrong firmware version	Bricked device	Versioning/control lapse	9/2/6/108	Automated version check, CRC verification, release-controlled builds	Require signed firmware and locked release pipeline, block manual override, log hash/version to device ID for traceability	SW-D4	27
Functional Test	Temperature drift 0.5°C	Comfort > complaint s/returns	Calibration error, sensor drift	9/2/6/108	Calibration routine, FT limits,	Add Cp/Cpk monitoring for calibration CTQ, introduce drift	QE-D7	42

periodic verification alarms and recalibration triggers, tighten supplier verification for sensors

Notes: rRPN = residual Risk Priority Number (post-action). Owner codes denote responsible role (ME, PE, VisEng, SW, QE) plus internal assignment ID. Here, ME – Manufacturing Engineer, PE – Process Engineer, VisEng – Vision/Inspection Engineer, SW – Software Engineer, and QE – Quality Engineer. Table 5 reports an illustrative extract emphasizing governance metadata; the step list is not intended to be a one-to-one match with Table 4.

In addition to drafting failure modes and controls, the GPT-5 output attached responsible-owner tags (e.g., ME-D3) and introduced rRPN to reflect expected risk reduction after the recommended actions. It also encoded decision logic tied to the predefined thresholds, including escalation/containment triggers (e.g., initiating a line hold when DPMO exceeds the θ_q limit). The linked control plan further included structured pass/fail rules and poka-yoke checks, such as automated verification of firmware CRCs and torque threshold compliance. Finally, KPI views and evidence bundles were organized into an audit-ready package to support faster review and traceability verification.

5.4. Comparative Evaluation

To compare the two NPI quality-planning solutions, we applied a structured set of validation indices that translate the framework's intended outcomes into observable evaluation criteria. The indices cover: 1) documentation completeness and audit readiness, 2) prevention-oriented detection and containment logic, 3) stability of capability controls, 4) governance and provenance safeguards, and 5) execution efficiency. They were informed by established quality and audit practice, NIST's definitions and guidance for trustworthy generative AI, and prior Quality 4.0 and industrial QA literature. Each index was scored independently by three QA experts, and the reported value is the mean of the three ratings.

Documentation Quality (DQ) reflects how complete, internally consistent, and audit-ready the deliverables are (e.g., cross-references, clause-to-evidence linkage, and consistent metadata). GPT-5 produced more explicitly traceable documentation packs, whereas the expert team delivered strong technical drafts with less formalized evidence mapping.

Detection and Containment (DC) captures how well the planning artifacts anticipate and manage nonconformities through clear containment rules, explicit "hold/stop" triggers, and coverage of rare or edge-case scenarios. The generative-AI condition scored higher due to more systematic inclusion of synthetic/low-frequency failure scenarios and threshold-linked containment logic.

Process Capability Stability (PCS) evaluates whether capability requirements on CTQs are not only stated but also supported with monitoring and reaction logic (e.g., trend alerts and degradation thresholds) that help sustain Cp/Cpk targets during ramp-up. Both approaches maintained Cp/Cpk ≥ 1.33 , while GPT-5 added earlier-warning mechanisms.

Governance and Provenance (GP) assesses audit defensibility through accountability controls such as role ownership, review/approval checkpoints, version control, and change traceability. The GPT-5 output incorporated clearer ownership tagging and change-trace structures; the expert workflow required more manual consolidation to achieve the same level of provenance formality.

Time Efficiency (TEI) measures effort reduction in producing the required artifacts and evidence package within the timebox, including automation of drafting, packaging, and trigger logic that otherwise increases coordination overhead. GPT-5 benefited from faster generation of audit-support elements, while the expert team relied more on manual integration.

Table 6 summarizes the comparative results. Conceptually, the indices map to the framework's decision logic: DQ supports document control and audit readiness; DC and PCS reflect prevention/containment behavior aligned with θ_q ; GP aligns with audit defensibility under θ_a ; and TEI captures efficiency within the shared five-phase cycle.

Table 6. Mean expert ratings comparing the expert-team and GPT-5-Thinking solutions.

Index	Expert team	GPT-5-Thinking	Rationale
DQ	3.0	3.8	GPT-5 provides traceable, audit-ready packs with clause links
DC	2.6	3.0	GPT-5 includes anomaly synthesis, expert solution lacks pre-pilot hold logic
PCS	3.0	3.4	Both ensure $C_p/C_{pk} \geq 1.33$; GPT-5 adds degradation alarms
GP	2.9	3.9	GPT-5 supports versioning, ownership, change-trace
TEI	3.2	3.7	GPT-5 auto-generates audits, triggers; expert solution requires manual integration

Note: Scores range from 0.0 to 4.0 and may take real (non-integer) values.

As summarized in Table 6, the GPT-5-Thinking condition achieved higher mean scores across all indices, with the largest margins in GP and DQ, reflecting stronger ownership tagging, versioning logic, and evidence linkage. Improvements in DC and PCS indicate broader pre-pilot risk coverage and more explicit monitoring/trigger rules aligned with the framework's θ_q gate, while the higher TEI score reflects reduced manual effort in drafting and packaging deliverables within the same timebox. Overall, the results suggest that, when constrained by the framework's document-control and approval requirements, generative AI can enhance audit defensibility (supporting θ_a) and planning efficiency without altering the underlying governance model – i.e., it strengthens performance within the framework rather than replacing it.

5.5. Discussion

This PoC indicates that the proposed framework can operationalize generative AI support for structured industrial QA tasks while preserving control requirements typical of regulated QMS environments. The comparison suggests that generative AI is particularly effective in areas where planning quality depends on structured documentation, cross-referencing, and evidence assembly, whereas expert teams remain strong in pragmatic judgment and context-specific refinement.

A key contribution of the framework is that it ties generative AI support to a repeatable operating rhythm (the five-phase cycle) and to explicit acceptance gates (θ_q , θ_a , and the overall threshold θ). In this way, AI outputs are treated as controlled inputs to the quality system: they must be reviewable, attributable, and traceable, and they can trigger corrective actions (when θ_q is not satisfied) or escalation (when θ_a is not satisfied) through defined governance pathways.

From an implementation standpoint, the findings support a hybrid deployment model. Generative AI can accelerate first-draft creation of PFMEAs, control plans, PPAP templates, and audit evidence packs, while human experts remain responsible for approval, risk acceptance, and final release through the QMS. This division of labor is consistent with the framework's emphasis on provenance, accountability, and auditable decision-making.

For manufacturing stakeholders, the framework provides a practical way to reduce planning cycle time and improve evidence continuity across teams. For auditors and compliance coordinators, it strengthens audit readiness by encouraging systematic trace linkage, ownership marking, and structured packaging of conformance evidence. More broadly, the PoC supports the claim that generative AI can be embedded into QA workflows in a standards-aligned manner – provided that organizations enforce document control, access governance, and human approval as non-negotiable requirements.

6. Conclusions and Future Research

This study proposed a structured framework for integrating generative AI into industrial quality assurance across a standardized five-phase cycle (planning, pre-run preparation, in-process monitoring, post-run correction, and review) and three nested assessment levels (product-, process-, and operation-level assurance). The framework also delineates stakeholder roles across manufacturer- and auditor-facing activities, and operationalizes decision gates through process-quality and audit-assurance thresholds (θ_q and θ_a). In doing so, it connects AI-enabled capabilities – such as PFMEA and control-plan drafting, anomaly interpretation, evidence packaging, and audit-oriented documentation – to established quality objectives and compliance expectations.

The proof-of-concept NPI case (smart thermostat ST-200) showed that a generative-AI-assisted workflow, when constrained by document control, review/approval checkpoints, and traceability requirements, can generate technically plausible planning artifacts and produce more audit-ready documentation packs within the same time constraints as an expert team. The comparative evaluation indicated consistent gains for the generative-AI condition, particularly in documentation completeness, governance/provenance, and time efficiency, suggesting that the framework can improve audit defensibility and reduce planning overhead without removing human accountability. At the same time, expert judgement remains critical for validating assumptions, confirming rare failure logic, and ensuring that AI-generated recommendations remain context-appropriate, especially where safety, regulatory exposure, or process changes are involved.

Based on these findings, three practitioner-oriented recommendations follow. First, AI deployment in QA should be implemented through explicit policies that define allowable use, required evidence, and approval gates, rather than treated as an informal productivity aid. Second, organizations should build capability through targeted training that combines quality standards literacy (e.g., document control, CAPA discipline, traceability expectations) with practical prompt and validation skills for generative tools. Third, governance should be designed up front – access control, versioning, source grounding, and role-based sign-off – so that AI-generated content can be defended during internal reviews and external audits. For supplier-facing processes, the framework supports earlier alignment on PPAP expectations, more consistent submissions, and clearer traceability in supplier communications.

This work has limitations. The empirical validation was intentionally bounded to a single PoC task and one product context, and it evaluated readiness through planning artifacts and structured indices rather than full-scale, long-run shop-floor deployment. In addition, performance depends on the maturity of the surrounding quality infrastructure (eQMS discipline, change control, and data availability); organizations with fragmented systems or weak document governance may see reduced benefits or higher operational risk.

Future research will expand validation across multiple products, plants, and industries, including settings with stricter regulatory regimes and higher safety criticality. We also plan to test the framework in live production environments to assess real-time monitoring, escalation behavior, and sustained human–AI collaboration outcomes over longer horizons. Finally, the θ_q and θ_a gates will be extended toward data-driven, adaptive thresholding based on historical capability, risk profiles, and evolving compliance requirements, while preserving auditability through transparent rules, change logs, and controlled approval workflows.

Author Contributions: Conceptualization, G.I., T.Y., and Y.I.; methodology, G.I. and Y.I.; validation, Y.I.; formal analysis, T.Y.; resources, G.I., T.Y., and V.H.; writing—original draft preparation, G.I.; writing—review and editing, G.I., T.Y., and Y.I.; visualization, G.I. and Y.I.; supervision, Y.I.; project administration, G.I.; funding acquisition, G.I., T.Y., and V.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially supported by the Project BG16RFPR002-1.014-0013-C01 „Digitalization of Economy in Big Data Environment– Second Stage“ (DIGD2) financed by the “Research, Innovation and Digitalization for Smart Transformation” Program 2021-2027 and co-funded by the European Union.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data supporting the findings of this study are contained within the article and no additional datasets were generated or analyzed.

Acknowledgments: The authors thank the academic editor and anonymous reviewers for their insightful comments and suggestions.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

8D	Eight Disciplines Problem Solving
AOI	Automated Optical Inspection
APQP	Advanced Product Quality Planning
BOM	Bill of Materials
CAPA	Corrective/preventive actions
CoPQ	Cost of Poor Quality
Cp	Process's potential capability
Cpk	Process capability index
CRC	Cyclic Redundancy Check
CTQ	Critical-to-Quality
DC	Direct Current
DMAIC	Define-Measure-Analyse-Improve-Control
DOE	Design of Experiment
DPMO	Defects per Million Opportunities
DPU	Defects per Unit
FMEA	Failure Mode and Effects Analysis
FPY	First Pass Yield
FT	Functional Test
GRR	Gauge Repeatability & Reproducibility
IPC	Acceptability Standard for Electronic Devices
LIMS	Laboratory Information Management System
LQC	Lean Quality Control
LSS	Lean Six Sigma
MSA	Measurement System Analysis
NCR	Non-Conformance Report
NPI	New Product Introduction
OEE	Overall Equipment Effectiveness
OEM	Original Equipment Manufacturing
PDCA	Plan-Do-Check-Act
PFMEA	Process Failure Mode and Effects Analysis
PLM	Product Lifecycle Management
PMIC	Power Management Integrated Circuit
PPAP	Production Part Approval Process
PPM	Parts Per Million
QA	Quality Assurance
QM	Quality Management
QMS	Quality Management System
RCA	Root Cause Analysis
SMT	Surface-Mount Technology
SPC	Statistical Process Control
S/O/D/RPN	Severity / Occurrence / Detection / Risk Priority Number

SOPs	Standard Operating Procedures
SPI	Solder Paste Inspection
SQE	Supplier Quality Engineer
SQM	Supplier Quality Management
TQM	Total Quality Management
V&V	Verification & Validation
XAI	Explainable AI

References

1. Escobar, C.A.; McGovern, M.E.; Morales-Menendez, R. Quality 4.0: A review of big data challenges in manufacturing. *J. Intell. Manuf.* **2021**, *32*, 2319–2334. <https://doi.org/10.1007/s10845-021-01765-4>.
2. Mahin, M.; Kadasah, N.; Alsabban, A.; Albliwi, S. Exploring the landscape of Quality 4.0: A comprehensive review of its benefits, challenges, and critical success factors. *Prod. Manuf. Res.* **2024**, *12*, 2373739. <https://doi.org/10.1080/21693277.2024.2373739>.
3. Khinvasara, T.; Ness, S.; Shankar, A. Leveraging AI for enhanced quality assurance in medical device manufacturing. *Asian J. Res. Comput. Sci.* **2024**, *17*, 13–35. <https://doi.org/10.9734/ajrcos/2024/v17i6454>.
4. Li, Y.; Zhao, H.; Jiang, H.; Pan, Y.; Liu, Z.; Wu, Z.; Shu, P.; Tian, J.; Yang, T.; Xu, S. Large language models for manufacturing. *arXiv*. **2024**, *arXiv:2410.21418*. Available online: <https://arxiv.org/abs/2410.21418> (accessed on 1 January 2026).
5. Mokander, J.; Sheth, M.; Gersbro-Sundler, M.; Blomgren, P.; Floridi, L. Challenges and best practices in corporate AI governance: Lessons from the biopharmaceutical industry. *Front. Comput. Sci.* **2022**, *4*, 1068361. <https://doi.org/10.3389/fcomp.2022.1068361>.
6. Chhetri, K.B. Applications of artificial intelligence and machine learning in food quality control and safety assessment. *Food Eng. Rev.* **2024**, *16*, 1–21. <https://doi.org/10.1007/s12393-023-09363-1>.
7. National Institute of Standards and Technology (NIST). Artificial Intelligence Risk Management Framework (AI RMF 1.0) (NIST AI 100-1); U.S. Department of Commerce: Gaithersburg, MD, USA, 2023. <https://doi.org/10.6028/NIST.AI.100-1>.
8. ISO/IEC 42001:2023. Information Technology - Artificial Intelligence - Management System; International Organization for Standardization and International Electrotechnical Commission: Geneva, Switzerland, 2023. Available online: <https://www.iso.org/standard/81230.html> (accessed on 1 January 2026).
9. European Union. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). *Off. J. Eur. Union* **2024**, *OJ L*, 2024/1689, 12.7.2024. Available online: <https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng> (accessed on 1 January 2026).
10. Cassoli, B.B.; Jourdan, N.; Nguyen, P.H.; Sen, S.; Garcia-Ceja, E.; Metternich, J. Frameworks for data-driven quality management in cyber-physical systems for manufacturing: A systematic review. *Procedia CIRP* **2022**, *112*, 567–572. <https://doi.org/10.1016/j.procir.2022.09.062>.
11. Ejjami, R.; Boussalham, K. Industry 5.0 in manufacturing: Enhancing resilience and responsibility through AI-driven predictive maintenance, quality control, and supply chain optimization. *Int. J. Multidiscip. Res.* **2024**, *6*. Available online: <https://pdfs.semanticscholar.org/3e49/9a40bdd3e7773ad93e2a7dcd1fdf4fabb6bf.pdf>.
12. NIST AI. Artificial intelligence risk management framework: Generative artificial intelligence profile (NIST AI 600-1); National Institute of Standards and Technology: Gaithersburg, MD, USA, 2024. <https://doi.org/10.6028/NIST.AI.600-1>.
13. Verma, H.; Padh, K.; Thelisson, E. *Can. AI be. Audi. arXiv*. **2025**, *arXiv:2509.00575*. Available online: <https://arxiv.org/abs/2509.00575> (accessed on 1 January 2026).
14. Raj Mohan, R.; Thirupathi, K.; Venkatraman, R.; Raghuraman, S. Quality improvement through first pass yield using statistical process control approach. *J. Appl. Sci.* **2012**, *12*, 985–991. <https://doi.org/10.3923/jas.2012.985.991>.
15. Verna, E.; Genta, G.; Galetto, M.; Franceschini, F. Defects-per-unit control chart for assembled products based on defect prediction models. *Int. J. Adv. Manuf. Technol.* **2022**, *119*, 2835–2846. <https://doi.org/10.1007/s00170-021-08157-1>.

16. Coskun, A.; Serteser, M.; Unsal, I. Sigma metric revisited: True known mistakes. *Biochem. Med. (Zagreb)* **2019**, *29*, 010902. <https://doi.org/10.11613/BM.2019.010902>.
17. Steiner, S.H.; MacKay, R.J. Effective monitoring of processes with parts per million defective: A hard problem! In *Frontiers in Statistical Quality Control 7*; Lenz, H.J.; Wilrich, P., Eds.; Physica-Verlag: Heidelberg, Germany, 2004; pp. 140-149. https://doi.org/10.1007/978-3-7908-2674-6_10.
18. ISO 22514-2:2017. Statistical methods in process management - Capability and performance - Part 2: Process capability and performance measures; International Organization for Standardization: Geneva, Switzerland, 2017. Available online: <https://www.iso.org/standard/71617.html> (accessed on 31 August 2025).
19. Muchiri, P.; Pintelon, L. Performance measurement using overall equipment effectiveness (OEE): Literature review and practical application discussion. *Int. J. Prod. Res.* **2008**, *46*, 3517–3535. <https://doi.org/10.1080/00207540601142645>.
20. Schiffauerova, A.; Thomson, V. A review of research on cost of quality models and best practices. *Int. J. Qual. Reliab. Manag.* **2006**, *23*, 647–669. <https://doi.org/10.1108/02656710610672470>.
21. Isack, H.D.; Mutingi, M.; Kandjeke, H.; Vashishth, A.; Chakraborty, A. Exploring the adoption of Lean principles in medical laboratory industry: Empirical evidences from Namibia. *Int. J. Lean Six Sigma* **2018**, *9*, 133–155. <https://doi.org/10.1108/IJLSS-02-2017-0017>.
22. Gueorguiev, T. An approach to integrate artificial intelligence in ISO 9001-based quality management systems. *Meas. Sens.* **2025**, *101787*. <https://doi.org/10.1016/j.measen.2024.101787>.
23. Sood, A.C.; Dhull, K.S. The future of Six Sigma - Integrating AI for continuous improvement. *Int. J. Innov. Res. Eng. Manag.* **2024**, *11*, 8–15. <https://doi.org/10.55524/ijirem.2024.11.5.2>.
24. Abdullah, R.; Lee, H.K.; Abdul Rasib, A.H.; Mansoor, H.O. Lean Six Sigma framework to improve the assembly process at a printer manufacturing company. *J. Adv. Manuf. Technol.* **2023**, *17*, 33–46. Available online: <https://jamt.utem.edu.my/jamt/article/download/6462/4004>.
25. Tanuska, P.; Spendra, L.; Kebisek, M.; Duris, R.; Stremy, M. Smart anomaly detection and prediction for assembly process maintenance in compliance with Industry 4.0. *Sensors* **2021**, *21*, 2376. <https://doi.org/10.3390/s21072376>.
26. El Hassani, I.; Masrour, T.; Kourouma, N.; Tavcar, J. AI-driven FMEA: Integration of large language models for faster and more accurate risk analysis. *Des. Sci.* **2025**, *11*, e10. <https://doi.org/10.1017/dsj.2025.7>.
27. Liu, H.C.; Liu, R.; Gu, X.; Yang, M. From total quality management to Quality 4.0: A systematic literature review and future research agenda. *Front. Eng. Manag.* **2023**, *10*, 191–205. <https://doi.org/10.1007/s42524-022-0243-z>.
28. ISO 9001:2015. Quality management systems - Requirements; International Organization for Standardization: Geneva, Switzerland, 2015. Available online: <https://www.iso.org/standard/62085.html> (accessed on 1 January 2026).
29. IATF 16949:2016. Quality management system requirements for automotive production and relevant service parts organizations; International Automotive Task Force: Southfield, MI, USA, 2016. Available online: <https://www.iatfglobaloversight.org/iatf-169492016/about> (accessed on 1 January 2026).
30. Doshi, J.A.; Desai, D. Overview of automotive core tools: Applications and benefits. *J. Inst. Eng. India Ser. C* **2017**, *98*, 515–526. <https://doi.org/10.1007/s40032-016-0288-z>.
31. ISO/IEC 17025:2017. General requirements for the competence of testing and calibration laboratories; International Organization for Standardization and International Electrotechnical Commission: Geneva, Switzerland, 2017. Available online: <https://www.iso.org/standard/66912.html> (accessed on 1 January 2026).
32. ISO 10012:2003. Measurement management systems - Requirements for measurement processes and measuring equipment; International Organization for Standardization: Geneva, Switzerland, 2003. Available online: <https://www.iso.org/standard/26033.html> (accessed on 1 January 2026).
33. ISO 19011:2018. Guidelines for auditing management systems; International Organization for Standardization: Geneva, Switzerland, 2018. Available online: <https://www.iso.org/standard/70017.html> (accessed on 1 January 2026).

34. Arunagiri, T.; Kannaiah, K.P.; Vasanthan, M. Enhancing pharmaceutical product quality with a comprehensive corrective and preventive actions (CAPA) framework: From reactive to proactive. *Cureus* **2024**, *16*, e69762. <https://doi.org/10.7759/cureus.69762>.
35. Holloway, S. The role of natural language processing in streamlining supply chain communication. *Prep.* **2024**, *202411.2303.v1*. <https://doi.org/10.20944/preprints202411.2303.v1>.
36. Scarton, G.; Formentini, M.; Romano, P. Automating quality control through an expert system. *Electron. Mark.* **2025**, *35*, 14. <https://doi.org/10.1007/s12525-025-00766-y>.
37. Friday, S.C.; Lawal, C.I.; Ayodeji, D.C.; Sobowale, A. Reviewing the effectiveness of digital audit tools in enhancing corporate transparency. *Int. J. Adv. Multidiscip. Res. Stud.* **2024**, *6*, 1679–1689. <https://doi.org/10.62225/2583049X.2024.4.6.4099>.
38. Kluska, R.A.; Rocha Loures, E.; Deschamps, F.; Camilotti, L.; Zanetti Freire, R.; Rotondo, R. Intelligent dashboard for asset management and maintenance with generative AI: A case study in maintenance engineering. In *Intelligent Production and Industry 5.0 with Human Touch, Resilience, and Circular Economy; Lecture Notes in Production Engineering; Sormaz, D.N.; Bidanda, B.; Alhawari, O.; Geng, Z., Eds.; Springer: Cham, Switzerland, 2025*. https://doi.org/10.1007/978-3-031-77723-3_60.
39. Moosavi, S.; Farajzadeh-Zanjani, M.; Razavi-Far, R.; Palade, V.; Saif, M. Explainable AI in manufacturing and industrial cyber-physical systems: A survey. *Electronics* **2024**, *13*, 3497. <https://doi.org/10.3390/electronics13173497>.
40. Tercan, H.; Meisen, T. Machine learning and deep learning based predictive quality in manufacturing: A systematic review. *J. Intell. Manuf.* **2022**, *33*, 1879–1905. <https://doi.org/10.1007/s10845-022-01963-8>.
41. Sun, Y.; Zhang, Q.; Bao, J.; Lu, Y.; Liu, S. Empowering digital twins with large language models for global temporal feature learning. *J. Manuf. Syst.* **2024**, *74*, 83–99. <https://doi.org/10.1016/j.jmsy.2024.02.015>.
42. Gao, R.X.; Kruger, J.; Merklein, M.; Kitzig-Frank, H.; Vancza, J. Artificial intelligence in manufacturing: State of the art, perspectives, and future directions. *CIRP Ann. - Manuf. Technol.* **2024**, *73*, 723–749. <https://doi.org/10.1016/j.cirp.2024.04.101>.
43. Liu, J.; Xie, G.; Wang, J.; Li, S.; Wang, C.; Zheng, F.; Jin, Y. Deep industrial image anomaly detection: A survey. *Mach. Intell. Res.* **2024**, *21*, 104–135. <https://doi.org/10.1007/s11633-023-1459-z>.
44. Xu, H.; Xu, S.; Yang, W. Unsupervised industrial anomaly detection with diffusion models. *J. Vis. Commun. Image Represent.* **2023**, *97*, 103983. <https://doi.org/10.1016/j.jvcir.2023.103983>.
45. Wan, Y.; Chen, Z.; Liu, Y.; Chen, C.; Packianather, M. Empowering LLMs by hybrid retrieval-augmented generation for domain-centric Q&A in smart manufacturing. *Adv. Eng. Inform.* **2025**, *65*, 103212. <https://doi.org/10.1016/j.aei.2025.103212>.
46. Jans, M.; Alles, M.G.; Vasarhelyi, M.A. A field study on the use of process mining of event logs as an analytical procedure in auditing. *Account. Rev.* **2014**, *89*, 1751–1773. <https://doi.org/10.2308/accr-50807>.
47. Rab, S.; Wan, M.; Sharma, R.K.; Kumar, L.; Zafer, A.; Saeed, K.; Yadav, S. Digital avatar of metrology. *Mapan* **2023**, *38*, 561–568. <https://doi.org/10.1007/s12647-023-00641-1>.
48. Picard, C.; Cozot, R.; Boissieux, L.; Segovia, B. Evaluating vision-language models for engineering design (including manufacturing and inspection tasks). *Artificial Intelligence Review* **2025**. <https://doi.org/10.1007/s10462-025-11290-y>.
49. Rydzy, S.; Zahradnikova, B.; Sutova, Z.; Ravas, M.; Hornacek, D.; Tanuska, P. A predictive quality inspection framework for the manufacturing process in the context of Industry 4.0. *Sensors* **2024**, *24*, 5644. <https://doi.org/10.3390/s24175644>.
50. Nguyen, H.T.T.; Nguyen, L.P.T.; Cao, H. XEdgeAI: A human-centered industrial inspection framework with data-centric explainable edge AI approach. *Inf. Fusion* **2025**, *116*, 102782. <https://doi.org/10.1016/j.inffus.2024.102782>.
51. Lin, Y.Z.; Shi, Q.; Yang, Z.; Latibari, B.S.; Shao, S.; Salehi, S.; Satam, P. Ddd-gendt: Dynamic data-driven generative digital twin framework. *IEEE Transactions on Artificial Intelligence* **2025**. <https://doi.ieeecomputersociety.org/10.1109/TAI.2025.3612920>.
52. Shafiee, S. Generative AI in manufacturing: A literature review of recent applications and future prospects. *Procedia CIRP* **2025**, *132*, 1–6. <https://doi.org/10.1016/j.procir.2025.01.001>.

53. Thomas, D. Revolutionizing failure modes and effects analysis with ChatGPT: Unleashing the power of AI language models. *J. Fail. Anal. Prev.* **2023**, *23*, 911–913. <https://doi.org/10.1007/s11668-023-01659-y>.
54. Alsaif, K.M.; Albeshri, A.A.; Khemakhem, M.A.; Eassa, F.E. Multimodal large language model-based fault detection and diagnosis in context of Industry 4.0. *Electronics* **2024**, *13*, 4912. <https://doi.org/10.3390/electronics13244912>.
55. Wang, T.; Zhang, B.; Jiang, D.; Li, D. A multimodal large language model framework for intelligent perception and decision-making in smart manufacturing. *Sensors* **2025**, *25*, 3072. <https://doi.org/10.3390/s25103072>.
56. Shadid, N.; Hamad, K.; Andi, O.; Alamassi, I. Revolutionizing quality management: Exploring Quality 4.0 frameworks in the context of Industry 4.0. In *Big Data in Finance: Transforming the Financial Landscape*, Volume 2; 2025; pp. 177-192. https://doi.org/10.1007/978-3-031-80656-8_16.
57. Alvaro, J.A.H.; Gonzalez Barreda, J. An advanced retrieval-augmented generation system for manufacturing quality control. *Adv. Eng. Inform.* **2025**, *64*, 103007. <https://doi.org/10.1016/j.aei.2024.103007>.
58. Wan, Y.; Chen, Z.; Liu, Y.; Packianather, M.; Wang, R. Empowering LLMs by hybrid retrieval-augmented generation for domain-centric Q&A in smart manufacturing. *Adv. Eng. Inform.* **2025**, *65*, 103212. <https://doi.org/10.1016/j.aei.2025.103212>.
59. Sun, Y.; Zhang, Q.; Bao, J.; Lu, Y.; Liu, S. Empowering digital twins with large language models for global temporal feature learning. *J. Manuf. Syst.* **2024**, *74*, 83–99. <https://doi.org/10.1016/j.jmsy.2024.02.015>.
60. Mata, O.; Ponce, P.; Perez, C.; Ramirez, M.; Anthony, B.; Russel, B.; Apte, P.; MacCleery, B.; Molina, A. Digital twin designs with generative AI: Crafting a comprehensive framework for manufacturing systems. *Journal of Intelligent Manufacturing* **2025**. <https://doi.org/10.1007/s10845-025-02583-8>.
61. ISO 22514-2:2017. Process capability and performance of time-dependent process models; International Organization for Standardization: Geneva, Switzerland, 2018. Available online: <https://www.iso.org/standard/71617.html> (accessed on 1 January 2026).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.