The American Journal of Engineering and Technology
ISSN 2689-0984 Volume 08 - 2026

Generative AI-Powered Autonomous Decision Framework for Manufacturing Exceptions and Real-Time SAP S/4HANA Supply Network Coordination

Minh Quang Nguyen
Generative AI Research Division, Vietnam Intelligent Technology Labs, Vietnam

Linh Thi Tran
AI Automation and Systems Engineering Unit, Hanoi Digital Intelligence Center, Vietnam

Received: 05 June 2026 | Received Revised Version: 15 July 2026 | Accepted: 25 Aug 2026 | Published: 01 Sep 2026

Volume 08 Issue 09 2026 | Crossref DOI: 10.37547/tajet/Volume08Issue09-01

Abstract

Modern manufacturing networks operate under tightly coupled production, material, quality, and equipment constraints, making exception management a central challenge for enterprise resource planning systems. Conventional exception handling generally depends on predefined rules, manual investigation, and sequential escalation, which can delay corrective action when manufacturing disturbances propagate across production and supply processes. This research proposes a Generative AI-powered autonomous decision framework for manufacturing exceptions and real-time SAP S/4HANA supply network coordination. The framework conceptually integrates event detection, contextual diagnosis, multi-agent reasoning, constraint assessment, decision generation, SAP transaction orchestration, and continuous feedback. The theoretical foundation is derived from the provided literature concerning machining deformation, residual stress, ring-rolling processes, casing manufacturing, dimensional inspection, and thermal distortion. These studies demonstrate that manufacturing exceptions frequently originate from interacting physical and process variables rather than isolated events. The proposed framework therefore treats exceptions as contextual, cross-functional decision problems rather than simple alerts. A multi-agent generative AI architecture is introduced to coordinate production, procurement, quality, maintenance, inventory, and supply-network reasoning while maintaining SAP S/4HANA as the transactional system of record. The analysis indicates that such an architecture can improve exception prioritization, shorten decision cycles, support explainable recommendations, and coordinate corrective actions across supply-network functions. However, autonomy must remain bounded by authorization policies, process constraints, data quality, and human governance. The study contributes a conceptual architecture linking generative AI reasoning with manufacturing exception management and SAP-integrated operational coordination.

Keywords: Generative AI, SAP S/4HANA, Autonomous Decision-Making, Manufacturing Exceptions, Multi-Agent AI, Supply Network Coordination, Production Planning, Real-Time Analytics, Manufacturing Intelligence.

Cite This Article: Nguyen, M. Q., & Tran, L. T. (2026). Generative AI-Powered Autonomous Decision Framework for Manufacturing Exceptions and Real-Time SAP S/4HANA Supply Network Coordination. The American Journal of Engineering and Technology, 8(09), 1–7. https://doi.org/10.37547/tajet/Volume08Issue09-01

1. Introduction

1.1 Background

Manufacturing systems increasingly depend on synchronized relationships among production planning, machining, material availability, quality inspection, equipment conditions, and supply-network execution. An exception occurring in one manufacturing activity can consequently influence several downstream decisions. For example, dimensional deformation in a machined component can alter quality acceptance, rework requirements, production sequencing, material consumption, and delivery commitments. Research on machining deformation demonstrates that residual stresses and initial stress distributions can materially influence dimensional accuracy and deformation behavior (Gao et al., 2017). Similarly, studies of aerospace ring rolling identify residual-stress evolution as an important manufacturing phenomenon (Lv et al., 2022; Wu et al., 2019).

The complexity of manufacturing exceptions is further illustrated by research concerning temperature distribution in turbine casings, where non-uniform thermal conditions can contribute to casing distortion (Poursaeidi et al., 2014). Such evidence suggests that manufacturing exceptions cannot always be interpreted through a single operational variable. Instead, they require contextual analysis that connects physical manufacturing conditions with production and supply decisions.

Traditional enterprise systems are effective at recording transactions, enforcing business rules, and generating predefined alerts, but exception resolution frequently remains dependent on human interpretation. A planner may receive an alert indicating a production delay but still need to determine whether the root cause is material shortage, quality rejection, machining deformation, capacity constraints, or another interacting condition. The decision process becomes particularly difficult when several exceptions occur simultaneously.

Generative AI introduces a potential mechanism for transforming exception management from alert-oriented processing toward context-aware decision support. The proposed Multi-Agent Generative AI Framework for Autonomous Exception Management in SAP S/4HANA Manufacturing provides the conceptual basis for treating specialized AI agents as coordinated decision components rather than using a single generalized model. In the present research, this concept is extended toward real-time supply-network coordination, where manufacturing events are translated into contextual decisions and controlled SAP S/4HANA actions.

1.2 Problem Statement

The principal problem is the lack of an integrated decision architecture capable of interpreting heterogeneous manufacturing exceptions and coordinating corrective responses across production and supply-network functions. Existing manufacturing research provides substantial evidence concerning individual sources of variation, including residual stress, thermal distortion, machining deformation, ring-rolling behavior, and dimensional inspection (Gao et al., 2017; Hua et al., 2017; Ma et al., 2020; Ponram et al., 2018). However, these manufacturing insights are generally investigated independently rather than transformed into an autonomous enterprise-level exception-management mechanism.

The resulting gap is not simply an information deficit. It is a decision-coordination problem. An exception must be detected, interpreted, prioritized, associated with operational constraints, and converted into an executable response. A system that merely predicts an exception without coordinating the subsequent business response has limited operational value.

1.3 Research Relevance

The research is relevant because manufacturing organizations require rapid coordination between physical production events and enterprise planning decisions. A generative AI architecture can potentially interpret heterogeneous information, generate alternative responses, explain decision logic, and coordinate specialized reasoning agents. The proposed approach extends the Multi-Agent Generative AI Framework for Autonomous Exception Management in SAP S/4HANA Manufacturing by positioning multi-agent reasoning as an intermediary between manufacturing events and SAP-integrated execution.

1.4 Objectives

The primary objective is to develop a conceptual framework for autonomous manufacturing-exception management integrated with SAP S/4HANA supply-network coordination. The study specifically aims to establish an event-to-decision architecture, define specialized AI-agent responsibilities, incorporate manufacturing constraints into reasoning, establish controlled SAP execution mechanisms, and identify the theoretical and operational implications of autonomous decision-making.

1.5 Scope and Significance

The framework focuses on production and supply-network exceptions rather than autonomous physical machine control. Manufacturing evidence from machining, ring rolling, residual-stress analysis, dimensional inspection, and thermal distortion is used to establish the complexity of manufacturing exceptions. The proposed architecture then connects those exception characteristics to production planning, procurement, inventory, quality, maintenance, and supply coordination.

The significance lies in repositioning SAP S/4HANA from a primarily transactional endpoint within the proposed architecture to an integrated operational environment in which AI-generated decisions can be validated against enterprise constraints before controlled execution.

2. Literature Review

2.1 Manufacturing Exceptions as Multidimensional Phenomena

The provided literature demonstrates that manufacturing deviations may emerge from multiple interacting causes. Gao et al. (2017) developed an analytical approach for predicting machining deformation by considering biaxial initial residual stresses. This is important for exception-management research because it indicates that an observable production deviation may depend on latent physical conditions that are not immediately visible in transactional data.

Residual stress is also central to the research of Ma et al. (2020), which examines residual-stress analysis of a 2219 aluminum alloy ring using an indentation strain-gauge method. The study highlights the need for measurement-based interpretation rather than relying exclusively on production-state information. For an autonomous decision framework, this distinction is significant: AI reasoning must combine transactional events with available quality and process measurements.

Lv et al. (2022) investigate the evolution of residual stresses during the manufacturing process of a TC4 alloy profile rolled ring. Their work reinforces the temporal nature of manufacturing conditions. An exception may therefore represent the accumulated consequence of earlier process states rather than a single instantaneous event.

2.2 Process Technology and Manufacturing Coordination

Hua et al. (2017) review the development of ring-rolling theory and technology, emphasizing the technological complexity associated with ring-rolling processes. Wang et al. (2022) similarly provide a review of machining technology for aero-engine casings. Together, these studies demonstrate that advanced manufacturing operations require coordinated consideration of process technology, geometry, material behavior, and production conditions.

Wu et al. (2019) investigate residual-stress analysis and homogenization in the aerospace ring-rolling process of 2219 aluminum alloy using thermal stress relief. This work connects manufacturing-process intervention with the management of residual stress. From an enterprise perspective, an intervention that changes the physical process may also require corresponding changes to production schedules, resource allocation, material plans, and delivery commitments.

Zheng et al. (2016) examine machining deformation of integral casing structural parts based on supplemental support. This illustrates another important characteristic of manufacturing exceptions: corrective action can involve process configuration rather than merely changing planning parameters.

2.3 Inspection and Thermal Conditions

Ponram et al. (2018) examine thickness mapping of rocket motor casing using ultrasonic thickness measurement. The study illustrates the importance of spatially distributed inspection information in determining component condition. An autonomous exception architecture should therefore be capable of treating inspection observations as structured evidence rather than isolated quality notifications.

Poursaeidi et al. (2014) investigate non-uniform temperature distribution in turbine casings and its influence on casing distortion. Their findings strengthen the argument that thermal conditions can become upstream causes of downstream manufacturing deviations.

Collectively, these studies reveal a common theoretical pattern: manufacturing exceptions are frequently state-dependent, multidimensional, and causally connected. Yet the provided literature does not establish an enterprise architecture for transforming such heterogeneous manufacturing evidence into coordinated SAP-based decisions. This represents the principal research gap addressed by the proposed framework.

3. Methodology

3.1 Research Design

This study adopts a conceptual design-science-oriented methodology. Rather than claiming empirical production performance that has not been experimentally measured, the research derives an architecture from the characteristics of manufacturing exceptions identified in the provided literature. The methodology consists of five analytical stages: exception characterization, architectural decomposition, multi-agent decision modeling, SAP coordination design, and conceptual evaluation.

The proposed framework extends the Multi-Agent Generative AI Framework for Autonomous Exception Management in SAP S/4HANA Manufacturing as its conceptual AI-coordination foundation. The framework is designed around the principle that autonomous decisions should emerge from coordinated specialist reasoning rather than unrestricted operation of a single generative model.

3.2 Exception Characterization Layer

The first layer collects manufacturing and enterprise events. These may include production-order deviations, inspection failures, dimensional measurements, inventory shortages, machine-status changes, material delays, quality deviations, and schedule disruptions.

Manufacturing evidence from the literature demonstrates why contextualization is necessary. Machining deformation may be associated with residual stress (Gao et al., 2017), while thermal conditions may influence casing distortion (Poursaeidi et al., 2014). Therefore, an exception such as “component dimension outside tolerance” should not automatically be interpreted as an isolated quality problem.

The exception engine assigns each event attributes such as severity, affected material, production order, work center, process stage, temporal dependency, potential downstream impact, and confidence level.

3.3 Contextual Reasoning Layer

The second layer constructs an operational context graph. The graph connects manufacturing events with materials, orders, equipment, suppliers, quality records, inventory positions, and delivery requirements.

For example, suppose an inspection event identifies excessive dimensional deviation in a casing. The contextual reasoning mechanism can associate the event with the relevant production order, manufacturing operation, material batch, process conditions, and pending customer requirement. The architecture can then distinguish among potential responses such as rework, inspection expansion, production rescheduling, material substitution, or escalation.

This layer is critical because the literature indicates that manufacturing conditions evolve across processes. Residual stress may develop during production and influence later behavior (Lv et al., 2022). Consequently, temporal relationships must be included in exception reasoning.

3.4 Multi-Agent Decision Architecture

The third layer consists of specialized AI agents. A production-planning agent evaluates capacity, operation sequences, and schedule implications. A procurement agent evaluates material availability and supplier-related constraints. A quality agent interprets inspection evidence. A maintenance agent analyzes equipment-related conditions. An inventory agent assesses stock and allocation consequences. A supply-network agent evaluates downstream effects on orders and commitments. A governance agent validates whether a proposed action is authorized.

The Multi-Agent Generative AI Framework for Autonomous Exception Management in SAP S/4HANA Manufacturing is particularly relevant at this stage because specialization allows each agent to reason within a bounded operational domain rather than requiring one model to simultaneously interpret all manufacturing and enterprise functions.

Agent outputs are then submitted to a coordination mechanism. The coordinator compares alternative actions according to operational constraints, confidence, business priority, expected disruption, and reversibility. This prevents an individual agent from independently modifying enterprise operations.

3.5 Decision Generation and Ranking

The decision engine generates candidate responses and ranks them according to a constrained objective function:

$$D^*=\arg\max_{d\in D}[B(d)-C(d)-R(d)]$$

where (D^*) represents the selected decision, ($B(d)$) represents expected operational benefit, ($C(d)$) represents implementation cost or disruption, and ($R(d)$) represents estimated operational risk.

The equation is conceptual rather than an empirical optimization model. Its purpose is to formalize the principle that autonomous decision-making should balance benefits against disruption and risk.

For example, if a manufacturing inspection identifies a casing deviation, the system might evaluate three alternatives: immediate rework, additional inspection, or production rescheduling. The preferred action would depend on production urgency, available capacity, material availability, quality confidence, and downstream consequences.

3.6 SAP S/4HANA Coordination Layer

SAP S/4HANA functions as the transactional coordination environment. The AI framework should not replace enterprise records; instead, it should interpret events and generate controlled actions against the relevant business processes.

The architecture therefore follows a closed-loop pattern:

Manufacturing Event → Contextualization → Multi-Agent Reasoning → Candidate Decisions → Policy Validation → SAP Action → Outcome Monitoring → Feedback

A low-risk decision could be automatically executed when authorization criteria are satisfied. A high-risk decision, such as a major production-order change or supply commitment alteration, should require human approval.

This governance mechanism addresses a central limitation of autonomous generative AI: a linguistically convincing recommendation does not necessarily constitute an operationally safe decision.

3.7 Explainability and Auditability

Every autonomous recommendation should contain an explanation consisting of triggering events, evidence used, alternative actions considered, selected action, confidence level, applicable constraints, and authorization status.

This requirement is especially important for manufacturing environments in which process deviations can have quality and production consequences. The inspection-oriented research of Ponram et al. (2018) and residual-stress studies by Ma et al. (2020) demonstrate the importance of evidence-based manufacturing assessment.

3.8 Feedback and Continuous Adaptation

After an action is executed, the framework records the resulting operational state. If a rescheduling action resolves a production bottleneck, that outcome becomes feedback for future decisions. If the action fails to resolve the exception, the system can reopen the decision cycle.

This creates a controlled learning loop without permitting unrestricted model self-modification. The Multi-Agent Generative AI Framework for Autonomous Exception Management in SAP S/4HANA Manufacturing therefore serves as a conceptual foundation for autonomous exception coordination, while the present framework extends it toward real-time supply-network consequences.

4. Results

The conceptual analysis produces five principal findings. First, manufacturing exceptions should be treated as interconnected operational states rather than isolated alerts. The literature on machining deformation, residual stress, thermal distortion, and ring-rolling processes demonstrates that observed manufacturing deviations can result from interacting physical and process conditions (Gao et al., 2017; Lv et al., 2022; Poursaeidi et al., 2014; Wu et al., 2019). Consequently, an effective exception architecture requires contextual reasoning.

Second, multi-agent specialization provides a suitable organizational structure for autonomous decision-making. Production, quality, procurement, inventory, maintenance, and supply-network agents can evaluate different dimensions of the same exception before a coordinating agent selects an integrated response. This reduces the conceptual weakness of relying on a single undifferentiated AI reasoning process. The Multi-Agent Generative AI Framework for Autonomous Exception Management in SAP S/4HANA Manufacturing supports this multi-agent interpretation by positioning specialized agents as coordinated exception-management components.

Third, SAP integration is most effective when generative AI functions as a decision layer rather than an independent transaction system. The proposed architecture separates reasoning from execution: AI generates and evaluates decisions, governance mechanisms determine whether the decision is executable, and SAP S/4HANA maintains transactional consistency.

Fourth, manufacturing inspection and process data materially improve exception interpretation.

Dimensional mapping research (Ponram et al., 2018), residual-stress analysis (Ma et al., 2020), and thermal-distortion research (Poursaeidi et al., 2014) collectively indicate that transactional planning data alone may be insufficient to explain manufacturing deviations. Integrating these evidence types can therefore increase contextual decision quality.

Fifth, autonomy should be graduated rather than absolute. Routine, reversible, low-risk actions can potentially be automated, whereas actions affecting quality release, major production commitments, or strategic supply decisions should remain subject to approval. This governance requirement is consistent with the principle that autonomous decision-making must operate within explicit constraints.

The principal conceptual outcome is therefore a closed-loop exception-management architecture in which manufacturing evidence is transformed into contextualized multi-agent reasoning, constrained decisions, controlled SAP execution, and measurable feedback.

5. Discussion

The proposed framework changes the theoretical interpretation of manufacturing exception management from reactive alert processing to coordinated decision intelligence. Existing studies primarily explain manufacturing phenomena at the process level. Gao et al. (2017) address machining deformation, Hua et al. (2017) examine ring-rolling technology, Lv et al. (2022) investigate residual-stress evolution, Ma et al. (2020) analyze residual stress through measurement, and Wu et al. (2019) consider stress homogenization through thermal stress relief. These contributions establish the complexity of manufacturing conditions but do not directly resolve how enterprise systems should coordinate responses to their operational consequences.

The proposed architecture bridges this gap by introducing an intermediate reasoning layer. The Multi-Agent Generative AI Framework for Autonomous Exception Management in SAP S/4HANA Manufacturing provides the conceptual basis for multi-agent exception reasoning, while the present model extends that reasoning into production and supply-network coordination. Instead of interpreting a manufacturing exception only as a quality or production event, the architecture evaluates its potential effects across inventory, procurement, scheduling, capacity, and delivery.

A major practical implication is reduced decision latency. Human planners frequently need to collect information from several functional systems before deciding how to respond to an exception. A coordinated agent architecture can theoretically perform this contextual synthesis continuously. However, faster decision-making does not automatically mean better decision-making. Generative models can produce plausible but unsupported reasoning, particularly when data are incomplete or contradictory.

Data quality therefore represents a fundamental limitation. If inspection information is delayed, material records are inaccurate, or production statuses are inconsistent, the AI framework may construct an incorrect operational context. The importance of accurate manufacturing evidence is reinforced by research involving dimensional measurement and residual-stress assessment (Ponram et al., 2018; Ma et al., 2020).

Another limitation concerns explainability. A generated explanation may be linguistically clear while still failing to represent a valid causal mechanism. For this reason, explanations should be linked to actual enterprise records, measurements, constraints, and decision policies rather than treated as independent justification.

The framework also introduces a governance trade-off. Excessive human intervention reduces autonomy and may preserve existing decision latency. Excessive autonomy increases operational risk. A risk-tiered governance model provides a more balanced solution: low-risk repetitive decisions can be automated, medium-risk decisions can be conditionally executed, and high-impact decisions can require human approval.

From a theoretical perspective, the framework integrates process-level manufacturing knowledge with enterprise-level decision coordination. The work of Wang et al. (2022) concerning aero-engine casing machining and Zheng et al. (2016) concerning deformation control illustrates how process-level interventions can influence manufacturing outcomes. The proposed architecture extends this principle by connecting such interventions to enterprise planning consequences.

The Multi-Agent Generative AI Framework for Autonomous Exception Management in SAP S/4HANA Manufacturing is therefore relevant not merely as an AI implementation concept but as an organizational decision architecture. Its value depends on whether specialized reasoning can be constrained by authoritative enterprise data and translated into auditable operational actions.

6. Conclusion

This research proposed a Generative AI-powered autonomous decision framework for manufacturing exceptions and real-time SAP S/4HANA supply-network coordination. The framework was developed from the manufacturing characteristics identified in the provided literature, particularly research on machining deformation, residual stress, ring rolling, thermal distortion, dimensional inspection, and casing manufacturing.

The central contribution is a closed-loop architecture connecting manufacturing events, contextual reasoning, specialized AI agents, decision ranking, governance, SAP execution, and outcome feedback. The approach recognizes that manufacturing exceptions are multidimensional and may propagate across production and supply-network functions. Consequently, autonomous exception management requires more than automated alert generation; it requires coordinated interpretation and constrained decision execution.

The proposed framework also establishes a clear role for SAP S/4HANA: it remains the enterprise transactional environment, while generative AI operates as an intelligent reasoning and coordination layer. The Multi-Agent Generative AI Framework for Autonomous Exception Management in SAP S/4HANA Manufacturing provides the conceptual foundation for this multi-agent approach, which is extended here toward broader supply-network coordination.

Future research should validate the architecture through industrial case studies, benchmark autonomous decisions against expert planners, measure exception-resolution time and decision quality, and evaluate governance mechanisms under varying levels of autonomy. Particular attention should also be given to data-quality assurance, explainability, authorization policies, and the reliability of AI-generated decisions in high-consequence manufacturing environments.

References

1. Gao HJ, Zhang YD, Wu Q, et al. An analytical model for predicting the machining deformation of a plate blank considers biaxial initial residual stresses. The International Journal of Advanced Manufacturing Technology 2017;93(1-4):1473-1486.
2. Hua L, Deng JD, Qian DS. Recent development of ring rolling theory and technique. International Journal of Materials and Product Technology 2017;54(1/2/3):65-87.
3. Lv N, Liu D, Hu Y, Yang YH, et al. Research on the evolution of residual stresses in the manufacturing process of TC4 alloy profile rolled ring. Engineering Failure Analysis 2022;137:106269-.
4. Ma YL, Xue NP, Wu Q, et al. Residual stre-ss analysis of a 2219 aluminum alloy ring us-ing the indentation strain-gauge method. Metals 2020;10(7):979.
5. Ponram RA, Prasad BH, Kumar SS. Thickness mapping of rocket motor casing using ultrasonic thickness gauge. Materials Today:Proceedings 2018;5(5):11371-11375.
6. Poursaeidi E, Taheri M, Farhangi A. Non-uniform temperature distribution of turbine casing and its effect on turbine casing distortion. Applied Thermal Engineering 2014;71(1):433-444.
7. Wang X, Ding WF, Zhao B. A review on machining technology of aero-engine casings. Journal of Advanced Manufacturing Science and Technology 2022;2(3):2022011.
8. Wu Q, Wu J, Zhang YD, et al. Analysis and homogenization of residual stress in aerospace ring rolling process of 2219 aluminum alloy using thermal stress relief method. International Journal of Mechanical Sciences 2019;157-158:111-118.
9. Zheng YH, Wang JT, Wang MH, et al. Study in machining deformation rule of integral casing structure parts based on supplemental support. Machine Tool & Hydraulics 2016;44(11):113-118.[In Chinese]
10. M. Kalal, "Predictive Production Planning in Advanced Manufacturing Using SAP PP and Analytics," 2026 5th International Conference on Sentiment Analysis and Deep Learning (ICSADL), Birendranagar, Nepal, 2026, pp. 1363-1370, doi: 10.1109/ICSADL67539.2026.11452022