Integrating Generative AI into SAP S/4HANA and SAP Analytics Cloud: A Framework for Autonomous Business Intelligence and Real-Time Enterprise Analytics

Dr. Lutfun Nahar

Assistant Professor, School of Computing and Data Science, Xiamen University Malaysia, Malaysia

Abstract

Business intelligence has long depended on analysts to extract data, build reports and interpret results. Generative artificial intelligence (GenAI) now makes a more autonomous model possible, in which systems detect change, explain it and propose actions with limited human effort. This article proposes a framework for integrating GenAI into SAP S/4HANA and SAP Analytics Cloud (SAC) to deliver autonomous business intelligence and real-time enterprise analytics. The framework rests on five layers: a clean-core S/4HANA foundation, a real-time analytics layer in SAC, a retrieval-augmented GenAI layer, an agentic orchestration layer, and a cross-cutting layer for privacy, security and resilience. Drawing on eight studies covering clean core modernisation, retrieval-augmented generation, reasoning agents, in-memory databases, data privacy governance and SAP landscape protection, the article shows how each layer contributes to faster and more reliable decisions. It concludes that autonomy is safe only when it is bounded: GenAI agents should act on governed data, within explicit permissions, and under human oversight for high-impact decisions.

Keywords: generative AI; SAP S/4HANA; SAP Analytics Cloud; autonomous business intelligence; real-time analytics; retrieval-augmented generation; clean core; data privacy governance

DOI: 10.21590/ijtmh.2023090447

1. Introduction

Most enterprises still run business intelligence as a request-and-response service. A manager asks a question, an analyst builds a query or dashboard, and the answer arrives after the moment for action has passed. Even where dashboards refresh in real time, people must notice the change, work out its cause and decide what to do.

The next stage is autonomous business intelligence. In this model, the analytics environment continuously monitors live data, detects meaningful change, explains it in plain language and proposes or initiates a response. Humans remain accountable, but their role shifts from compiling information to judging recommendations.

SAP environments are well placed for this shift. SAP S/4HANA runs transactional and analytical workloads on one in-memory platform, so analysis can use the current state of the business rather than an overnight copy (Plattner, 2009). SAP Analytics Cloud adds live dashboards, integrated planning and predictive models on top of that data. Generative AI supplies the missing ingredient: an interface that can read, reason over and narrate enterprise data, and increasingly take actions through tools.

Integration, however, is not simple. Large ERP estates contain years of custom code that slows upgrades and blocks new capabilities. Language models can generate confident but wrong answers when they are not grounded in authoritative data. ERP systems also hold sensitive personal and financial data, and they are attractive targets for malware and supply-chain attacks. An autonomous system built on a weak foundation would act quickly on bad information.

This article addresses these challenges by proposing a layered framework. It asks three questions. What architectural foundations does autonomous BI require in an SAP landscape? How can GenAI be grounded and orchestrated so that its outputs are reliable? And what governance controls keep autonomy safe? Section 2 reviews the literature, Section 3 presents the framework, Section 4 illustrates applications, Section 5 addresses privacy, security and resilience, Section 6 sets out a roadmap and research agenda, and Section 7 concludes.

2. Literature review

The studies informing this framework fall into three groups: the SAP foundation, generative and agentic AI, and the governance of enterprise data and infrastructure.

2.1 The SAP foundation: real-time data and clean core

Plattner (2009) challenged the separation of transactional (OLTP) and analytical (OLAP) systems. He argued that an in-memory column store could serve both workloads from one dataset, removing pre-computed aggregates and batch extraction. This principle underpins S/4HANA and makes real-time analytics on operational data possible.

Asuri (2021) examines how modernising SAP supplier relationship management through a clean core architecture prepares organisations for the S/4HANA transition. Clean core means keeping the standard ERP free of invasive custom code and moving extensions to well-defined, upgrade-safe layers. For this article, the significance is direct. A clean core gives GenAI stable, standard data models and APIs to work with, rather than a tangle of bespoke tables and logic that an AI agent cannot interpret reliably. It also supports the business process excellence on which trustworthy analytics depend.

Sarferaz (2025) addresses how GenAI business applications can be systematically developed and operated within ERP software. Given the scale of ERP systems, which cover tens of thousands of business processes, the study argues for a structured lifecycle rather than isolated experiments. This reinforces the case for a framework-based approach.

2.2 Generative and agentic AI

Lewis et al. (2020) introduced retrieval-augmented generation (RAG), which pairs a language model with a retriever over an external document index. RAG produced more specific and factual output than generation alone, and its knowledge could be refreshed by updating the index rather than retraining the model. In an enterprise, the index can hold master data, analytical results and policy documents, so that answers reflect the organisation's own facts.

Cherukuri (2023) applies GenAI and RAG to software engineering, covering automated code generation and intelligent debugging. Two lessons transfer to SAP analytics. First, grounding generation in a relevant knowledge base improves the accuracy of technical output. Second, GenAI can accelerate the build work behind analytics, such as generating data models, queries, transformation logic and extension code, while debugging assistance reduces the time spent fixing them. This matters because the bottleneck in many SAP analytics programmes is development capacity, not ideas.

Yao et al. (2023) proposed ReAct, an approach in which language models interleave reasoning steps with actions such as tool calls. This lets a model gather information, adjust its plan and act in sequence. ReAct provides the conceptual basis for autonomous BI agents that query SAC, retrieve S/4HANA records and decide what to do next.

2.3 Governance of data and infrastructure

Asuri (2022) sets out an approach to enterprise data privacy governance across SAP ecosystems. It combines context-aware masking, enforcement of retention rules across the data lifecycle, and compliance with regulations across jurisdictions. These controls become critical once GenAI can read ERP data, since a single prompt could otherwise expose personal data to users or external models.

Asuri (2023) focuses on safeguarding mission-critical SAP landscapes. It covers advanced malware containment, assurance of third-party software, and resilient recovery mechanisms for hybrid infrastructure. Autonomous analytics increases dependence on SAP availability and integrity. It also introduces new third-party components, such as model providers and integration services, which must be assured in the same way.

The literature therefore supplies the components of autonomous BI, from real-time data and clean architecture to grounded generation, reasoning agents and governance. What is missing is an integrated framework that shows how these components fit together in an SAP landscape. Section 3 provides one.

3. The proposed framework

The framework has four stacked layers and one cross-cutting layer. Each stacked layer depends on the integrity of the one beneath it.

1. **Clean-core foundation (SAP S/4HANA).** S/4HANA is the single system of record, running transactions and analytics on live in-memory data (Plattner, 2009). Following clean core principles, custom logic is moved out of the core into upgrade-safe extensions, and data is exposed through standard, documented interfaces (Asuri, 2021). This gives every higher layer a consistent, well-understood view of the business.
2. **Real-time analytics layer (SAP Analytics Cloud).** SAC connects live to S/4HANA and other sources. It hosts key performance indicators, planning models and predictive functions such as forecasting and anomaly detection. In an autonomous design, these models run continuously and publish signals, such as a forecast breach or an unusual cost pattern, rather than waiting for a user to open a dashboard.
3. **Grounded generative layer.** Language models interpret signals and answer questions using RAG (Lewis et al., 2020). The retrieval index holds S/4HANA master and transactional data, SAC results, business glossaries and policies. Every answer is tied to retrievable sources, which reduces fabricated output and lets users verify claims. The same layer supports developers by generating and debugging analytics artefacts (Cherukuri, 2023).
4. **Agentic orchestration layer.** Agents follow a reason-and-act cycle (Yao et al., 2023). Triggered by a signal, an agent can query SAC for detail, retrieve the underlying documents, run a what-if scenario, draft an explanation and propose an action. Actions are classed by risk. Low-risk actions, such as sending an alert or refreshing a forecast, can run automatically. High-risk actions, such as changing a purchase order or a credit limit, require human approval before they are written back to S/4HANA.

The cross-cutting **trust layer** applies privacy governance, including masking, retention and jurisdictional rules (Asuri, 2022), together with landscape security, third-party assurance and recovery (Asuri, 2023). It operates at every layer, so that agents see only the data a user is entitled to see and the platform can recover if any component fails.

The framework embodies three principles. First, **grounding before generation**: no autonomous output without traceable sources. Second, **bounded autonomy**: the scope of automatic action is explicit and widened only as trust grows. Third, **lifecycle discipline**: GenAI features are developed, tested, monitored and retired through a managed process, as Sarferaz (2025) recommends for ERP systems.

4. Autonomous business intelligence in practice

The framework changes how four common analytics tasks are performed.

Continuous financial monitoring. Instead of waiting for month-end, SAC models track margin, working capital and cost variances daily against plan. When a threshold is breached, an agent drills into S/4HANA line items, identifies the drivers and drafts a short explanation for the controller. The controller reviews the cause, not the spreadsheet.

Supplier and procurement intelligence. A clean-core supplier relationship management process (Asuri, 2021) produces consistent supplier, contract and delivery data. Agents can monitor delivery performance and price trends, flag suppliers at risk and propose alternative sourcing scenarios. Buyers approve any change before it reaches the purchasing documents.

Conversational self-service analytics. Business users ask questions in natural language, such as which regions missed their sales targets last week and why. The RAG layer retrieves the relevant SAC figures and S/4HANA records and returns an answer with its sources. This widens access to analytics without multiplying custom reports.

Faster analytics delivery. When a new metric or data model is needed, GenAI can generate a first draft of the query, calculation logic or extension code and help debug it (Cherukuri, 2023). Developers review and test the output, shortening delivery cycles while keeping custom code outside the clean core.

In each case, the value comes from shortening the loop between signal, understanding and action. Real-time data removes the reporting delay, while grounded GenAI removes much of the interpretation delay. Bounded autonomy, in turn, reduces the execution delay without removing accountability.

5. Privacy, security and resilience

Autonomy raises the stakes of every weakness in the data platform, so the trust layer must be designed in from the start.

Privacy. Context-aware masking should apply before data enters a prompt or retrieval index, so that agents receive only what the user's role and purpose allow. Retention rules must extend to embeddings, caches and conversation logs, and cross-border data flows to model providers must satisfy each relevant jurisdiction (Asuri, 2022).

Security. GenAI connectors, model services and agent tools are new third-party components and should pass the same assurance checks as any SAP add-on. Agent permissions should follow least privilege, and every action should be logged. Malware containment and segmentation limit the damage if a component is compromised (Asuri, 2023).

Resilience. Hybrid recovery plans should cover the retrieval index, model configurations and agent workflows, not only the S/4HANA database (Asuri, 2023). If the AI layer fails, core transactions and standard SAC reporting must continue unaffected.

6. Implementation roadmap, challenges and future research

Adoption works best in three phases. First, establish the foundation: clean the core, standardise data models and put privacy controls in place. Second, introduce grounded conversational analytics on one or two well-governed domains. Third, add agents with narrowly defined, low-risk actions, and widen their scope as accuracy is demonstrated.

The main challenges are data quality, the cost of running models at enterprise scale, skills gaps, and resistance to delegating decisions to software. Future research should measure the effect of autonomous BI on decision speed and accuracy in real SAP deployments. It should also develop evaluation methods for agent behaviour and study how regulators will treat AI-initiated business actions.

7. Conclusion

Integrating GenAI into SAP S/4HANA and SAP Analytics Cloud can turn business intelligence from a reporting service into an autonomous capability that detects, explains and acts on change in real time. The proposed framework shows that this depends on a clean-core foundation, grounded generation, reasoning agents with bounded authority, and a trust layer covering privacy, security and resilience. Organisations that build these layers in sequence can gain speed without sacrificing control.

References

Asuri, R. (2021). Advancing enterprise digital transformation by modernizing SAP supplier relationship management through clean core architecture for SAP S/4HANA transition readiness and business process excellence. *International Journal of Recent Advances in Science and Technology, 8*(04), 1–20.

Asuri, R. (2022). Establishing enterprise data privacy governance across SAP ecosystems through context-aware masking, lifecycle retention enforcement, and cross-jurisdictional regulatory compliance. *International Journal of Humanities and Information Technology, 4*(01-03), 352–375.

Asuri, R. (2023). Safeguarding mission-critical SAP landscapes through advanced malware containment, third-party software assurance, and resilient hybrid infrastructure recovery mechanisms. *SAMRIDDHI: A Journal of Physical Sciences, Engineering and Technology, 15*(03), 475–493.

Cherukuri, M. (2023). Leveraging generative artificial intelligence and retrieval-augmented generation to improve software engineering, automated code generation, and intelligent debugging systems. *International Journal of Humanities and Information Technology, 5*(04), 140–163. https://doi.org/10.21590/ijhit.05.04.11

Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W., Rocktäschel, T., Riedel, S., & Kiela, D. (2020). Retrieval-augmented generation for knowledge-intensive NLP tasks. *Advances in Neural Information Processing Systems, 33*. https://arxiv.org/abs/2005.11401

Plattner, H. (2009). A common database approach for OLTP and OLAP using an in-memory column database. In *Proceedings of the 2009 ACM SIGMOD International Conference on Management of Data* (pp. 1–2). ACM. https://doi.org/10.1145/1559845.1559846

Sarferaz, S. (2025). Implementing generative AI into ERP software. *IEEE Access, 13*, 73342–73354. https://doi.org/10.1109/ACCESS.2025.3564133

Yao, S., Zhao, J., Yu, D., Du, N., Shafran, I., Narasimhan, K., & Cao, Y. (2023). ReAct: Synergizing reasoning and acting in language models. In *Proceedings of the International Conference on Learning Representations (ICLR 2023)*. https://arxiv.org/abs/2210.03629