A review on next generation ERP in era of AI and ML implementation frameworks

Bijendra Tyagi[1], Arjun Singh[2], Atharv Verma[3], Amit Rajawat[4], Tushar Chaudhary[5]

[1, 2,3,4,5] CSE Dept. JSS Academy of Technical Education Noida, UP, India

tyagig100@gmail.com[1], arjunsingh120415@gmail.com[2], atharv.verma29@gmail.com[3], kamitrajawat@gmail.com[4], chuadharytushar017@gmail.com[5]

Abstract

With the rapid evolution of AI and machine learning models, there is a growing need of using AI and ML in business decision making and other enterprise use cases. This paper proposes a comprehensive review of the most recent major research advancements in the field of AI and ML integration with Enterprise Resource Planning platforms to increase efficiency, it also explores the key drawbacks or limitations of the research methodology so far. This paper compiles a list of major research gaps and aims to fill some of them by presenting a new AI/ML powered ERP system architecture or framework based on microservice architecture for loose coupling, enhanced scalability, faster deployment times and increased agility. This architecture allows applications to run as small, independent services which provides flexibility in technology choices and features and makes it easier to be maintained by a small team of developers.

Keywords: *Artificial Intelligence (AI), Machine Learning (ML), Enterprise Resource Planning (ERP) Systems, Predictive Analytics, Generative AI, Microservices (MS), MLOps, Extract Transform Load (ETL) Pipelines, Real-Time Analytics, Large Language Models (LLMs)*

1. Introduction

Enterprise Resource Planning (ERP) systems are a central software suite which provide access to various business functions and processes, such as supply chain, human resources, manufacturing, etc. with a central database. The implementation of ERP has shown to massively increase productivity and efficiency in business operations and management.

With the rising improvement in AI and ML models and systems and how they have increased efficiency in other domains, enterprises and businesses also want to leverage AI and ML systems to increase efficiency in business management as well as faster decision-making. This paper focuses on reviewing existing research studies and their proposed methodologies regarding improvement in traditional ERP systems with modern software engineering practices [4, 16] and AI/ML integration with predictive analytics in real-time [1, 2, 7, 8].

The comprehensive review defines the major research gaps and problems that still blocks the development of a standardized AI-powered ERP system. Some of these research gaps are found to be regarding data security and privacy, running costs, and the general lack of a standardized framework that caters to every enterprise.

This paper proposes a new architecture that aims to fill these research gaps by addressing the data security and privacy concerns, cost optimization problem, and also introduces a standardized method that can be universally reusable

2. Literature Review

The objective of this section is to thematically review existing methodologies that have made major contributions to the next generation of ERP systems.

2.1 Architectural Bases for Next Generation ERP

The legacy, monolithic software systems have some well-known struggles with respect to increasing complexity of modern software applications. Monolithic software systems struggle to keep up with growth and the ever-changing nature of modern software system requirements and they struggle with agility, and flexibility due to tightly coupled modules that make it hard to add new features or change existing ones [4]. Microservices architectures (MSA) are considered as the modern standard for software applications because they provide scalability, maintainability, and readability of codebases. MSA uses loosely coupled components with high cohesion meaning every component is independent of other components, this allows programmers to easily add new modules or modify a module as there is only one module that has a particular function [4]. The field of Microservice Identification (MSI) is an evolving field with no universally accepted architecture or framework yet. As discussed in [4], this evolving research field is introducing various new tools and techniques to help programmers understand how they can transition from monolithic architecture to an MSA.

With everything moved online, all the data and businesses have also moved online. This new business and software needs required global low latency sharing of data and services so that anyone can access and are not location bound. Cloud computing meets these needs by hosting services and data on a global network so that anyone can access those globally [16]. Cloud computing has changed every business domain all around the world and made them really scalable and accessible [14, 17].

2.2 Advancements in Data handling & Analytics

With the arrival of machine learning and artificial intelligence, a stronger data handling framework has appeared as "Data Pipelines". Prior research [1], shows that Extract Transform and Load (ETL) pipelines are one of the best methods of data transformation in the modern AI era. They can perform all the major data transformations required for using the data in ML models by performing steps like missing data handling, and feature engineering on raw data and then store it in data warehouses for further model training or analysis. AI-powered ETL pipelines can take advantage of automation of AI to perform actions like automatic data handling, transforming the data automatically based on patterns in data, anomaly detection in data, and adapt to changes in data by adjusting the database schema [1].

Real-time Analytics (RTA) architectures are also pretty common nowadays that can handle high-velocity data and AI-driven analytics tools [7]. Increasing use of AI and ML in RTA has also made real-time decision making and automation possible as well. The use large language models (LLMs) in business analytics have also been a major advancement in predictive data analytics which can provide LLM based value assessment based on the data [20].

2.3 Operationalizing AI and ML models

Advancements in Natural Language Processing (NLP) techniques have allowed LLMs to be able to understand human language and made AI much more versatile in creation of AI-powered chatbots allowing for smart decision-making and efficiency [19]. The use of chatbots

have improved customer experience and business applications by simplification of mundane tasks and more powerful abilities that chatbots allow in various business use cases [15, 20, 21]. In production environments, simple or linear ML model frameworks cannot be used because they are unable to adapt to changes in data over time and may incur high maintenance costs down the line. Using DevOps practices of continuous development and integration (CI/CD) in ML frameworks can solve this problem and make production ready ML systems possible [3]. The process of applying CI/CD pipelines in machine learning systems can solve the problem of updating models over time and it could reduce maintenance costs as well. Inspired by DevOps, Machine Learning Operations also known as, MLOps has been introduced [2]. It provides a production ready framework that uses all the best practices of continuous retraining of ML models to stay relevant over time and continuous development and changes can reduce the risk of costly maintenance actions, which can improve operational efficiency in ML systems [2].

2.4 Application domains of AI and ML in ERP systems

The application of AI and ML is transforming the way businesses operate their enterprise systems to enable them to respond quicker and smoothly daily operations. They provide insight in real time, easier workflow and coordination of large processes [12]. This is further expanded by new concepts like agent-based designs and compound AI systems [6] which allow organizations to integrate their own data, APIs and models in a single set up. Most ERP systems become sluggish with an increase in business data, particularly those in finance. This is addressed in [5] by the method that introduces smart agents that learn what users want and create tasks as soon as they are needed. This saved almost 40 percent processing time and approximately 94 percent of significant errors in banking tests [5]. The SAP S/4HANA also facilitates this change by applying in-memory computing and built-in AI capabilities to generate reports more quickly, insights into the future, and automatically run financial processes and is therefore simpler to create a business that grows its performance without compromising performance [11].

Category	Legacy ERP	System / AI-Driven ERP
Architecture	Monolithic, hard-coded workflows	Modular, service-based, ML-augmented architecture
Response Time	Batch processing; slow insights	Real-time predictions and instant analytics
Flexibility	Low; fixed rules, difficult to update	High; adaptive models and dynamic rule systems
Scalability	Limited horizontal scaling; dependent on central servers	Cloud-native scaling; distributed microservices
Intelligence Level	Rule-based automation only	Predictive + prescriptive intelligence (ML, NLP, agents)
Integration Capability	Difficult API integrations	API-first, integrates with ML pipelines and cloud tools
Maintenance Cost	High manual maintenance; frequent downtime	Automated monitoring, self-learning pipelines reduce cost

User Experience	Static dashboards and reports	Interactive dashboards, personalized insights
Data Handling	Structured data only (SQL-heavy)	Structured + unstructured (text, logs, images)
Customization	Slow, vendor-dependent	Fast, model-driven, customizable workflows

Businesses are subjected to rapid changes and therefore supply chains must be able to respond more quickly to problems. The area of work in [8] presents an approach that predicts risks prior to their occurrence with the help of machine learning and predictive analytics rather than just relying on old data and responding to it later. The system learns patterns by training its models on the past trends and the current context which indicates a likely future issue. This arrangement coupled with real-time monitoring is useful by helping firms to prevent delays and act faster, and early outcomes demonstrate improved visibility, and more flexible supply chain [8].

2.5 Comparison of legacy ERP and AI driven ERP

Table 2.1: Comparison of legacy and AI driven ERP

Table 2.1: Comparative Analysis of Existing Research

Ref.	Author	Application Area	Key Findings	Research Gaps
[4]	I. Oumoussa et al., 2024	Microservices Architecture	Solves issues of concurrency and scalability in the registration system and reduced turn-around-time.	Result may not be replicated in a larger model, suboptimal use of resources in current implementation, lack of advanced ML techniques.

[7]	W. Chen et al., 2025	Analytics	The paper's main purpose is to provide a comprehensive guide and blueprint for building these Real-Time Analytics (RTA) solutions, specifically highlighting how to integrate advanced Machine Learning (ML) and Artificial Intelligence (AI) tools into the process.	Currently, the systems designed to spot patterns in fast data (Complex Event Processing, or CEP) and the smart systems that make predictions (Machine Learning, or ML) are seen as two distinct, separate areas. We need to figure out better ways to allow ML models to learn continuously in real-time from data streams (incremental learning).
[1]	S. Kandhikonda et al., 2025	ETL Pipelines	The use of AI changes the ETL process from mechanical work to an automated, self-optimizing process. Reduction of data quality issues with reduced turn-around-time along with compliance of classified data.	The process of reducing tech debt in a large organisation is cumbersome.
[2]	D. Kreuzberger et al., 2023	MLOps	Implementation of ML projects in real business environment with the help of CI/CD principles. The importance of different roles is also discussed to build and operate the ML infrastructure.	The shortage of skilled experts for key roles required of ML infrastructure. Adopting MLOps requires a major culture shift from the primitive approach. The need of collaboration and communication between different teams to build the system.

[5]	H. Yang et al., 2025	Finance /ERP	The paper introduces a significant framework called FinRobot, which uses Generative Business Process AI Agents (GBPAs) to fundamentally replace the inflexible, outdated workflows.	Complex implementation, outdated systems, rigid business requirements and complex
[14]	Bandara et al, 2023	ERP	This study investigated how integrating Big Data technologies with Enterprise Resource Planning (ERP) systems can make those systems more responsive to large amounts of data	Organizations are currently struggling to manage the integration between ERP systems and big data technologies. This struggle leads to significant inefficiency, as many organizations use only 12% of the collected data, leaving 88% wasted
[16]	V. Chang et al, 2015	Cloud Computing	The Cloud Computing Adoption Framework (CCAF), a useful, multi-layered security architecture intended to safeguard sensitive data and services stored in high-stakes enterprise clouds, is presented in this study. Testing proved the framework's effectiveness: it successfully detected and blocked malicious files, was resistant to continuous attacks, and showed an excellent ability not to report false alarms.	The primary limitation is that the viruses and trojans used to test the system were based on known vulnerabilities from 2013 because more current samples were unavailable for testing at the time of publication

[8]	A. Aljohani, 2023	Analytics	This paper focuses on how businesses can use Predictive Analytics and Machine Learning (ML) to fundamentally change how they manage their Supply Chains, making them faster and more adaptable.	Existing research often focused only on isolated pieces of risk management (e.g., just identifying a risk or just evaluating its impact). There was a clear gap in creating a seamless, holistic system that connects risk identification, impact evaluation, and proactive response strategies all together.
[18]	Jianxun Jiang et al, 2026	Deep Learning	The paper proposes an innovative strategy for Supply Chain Risk Management that uses Predictive Analytics and Machine Learning (ML) to transition management from reactive to proactive risk mitigation.	Previous research often focused on isolated parts of risk management (e.g., only identifying a risk or only assessing its impact). The major gap was the lack of a seamless, holistic system that connects risk identification, impact evaluation, and proactive response strategies all together.
[20]	W. De Michele *et al,* 2025	LLM	This research addresses the problem that most organizations optimize their operations through Business Process Management (BPM), but the crucial step of analysing those processes—especially identifying wasted effort—is currently slow, expensive, subjective, and requires manual effort from experts.	The current analysis relies on the general language understanding capabilities of the LLM and the process description, but it may miss domain-specific details or terminology (Data and Domain Dependence).

3. Proposed Methodology

This paper proposes "Insighto – Cognitive Business Suite", an attempt to resolve some of the drawbacks discussed in Section 3 by designing an adaptive AI workflow for a specific ERP module.

3.1 Overview

Insighto is a novel architecture that aims to provide uniform AI-powered data pipelines that can work for varying company data along with AI-powered insights and recommendations. Insighto attempts to allow business decisions to be made faster.

3.2 Key Features

1)   In this architecture, we propose a pipeline that at the beginning will be converting heterogenous forms of raw company data to a standard set of features that will allow the workflow to still be usable by varying enterprise naming conventions in their data [1]. Along with this initial conversion, we plan to use MLOps practices to reduce operational overhead [2, 3].

2)   As part of the CI/CD pipelines of MLOps, we will incorporate "Champion model & Challenger model" workflow with iterative retraining and challenging the production model in an attempt to ensure the production model is relevant and can adapt over time [2].

3)   The final component of this architecture is to use secure, self-hosted Large Language Models (LLMs) as agents which can ensure that sensitive data remains secure within the company servers. The tasks of the agents would be to take the statistical analysis of the data and model predictions along with core "eXplainable AI (XAI)" feature importance metrics mentioned in [9] and provide actionable business insights and recommendations as the output, which should increase transparency of the reasoning behind agent's recommendations [7].

3.3 System Architecture

Figure 3.1 AI and ML pipeline architecture
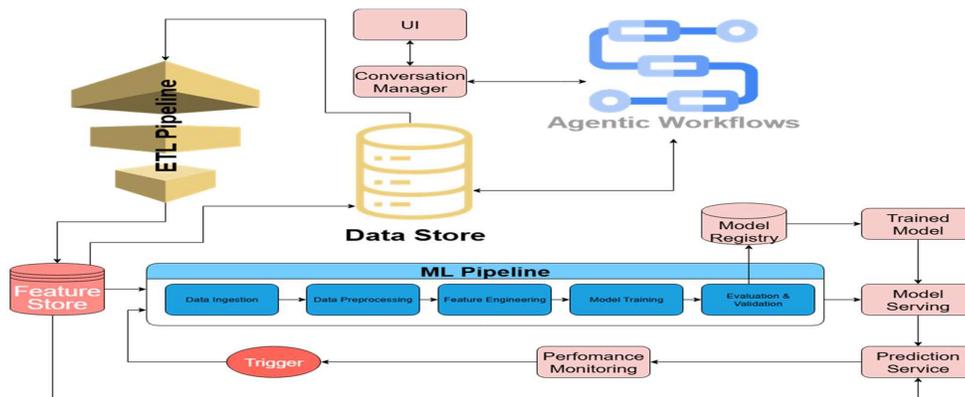
ERP Layer (Diagram 3.2)

• The system blends a Predictive ERP platform with an agentic and ML-driven workflow pipeline.

• User actions from web or mobile apps pass through an API Gateway for authentication, validation, and logging.

• Core microservices handle login processes, ERP data operations, and conversational features.

- AI components generate real-time predictions, forecasting outputs, and chatbot responses.

- Continuous logs from all services feed into a central monitoring setup, while cloud storage and a message broker support scalable and event-driven communication.

ML + Agentic Pipeline (Diagram 3.1)

- The ML pipeline manages the full development cycle of models used for ERP predictions and forecasting.

- ERP data moves through an ETL/ELT process into storage and a Feature Store consumed by the pipeline.

- Automated steps include data ingestion, preprocessing, feature creation, training, and evaluation.

-

- Training jobs are triggered by events or schedules, with model performance monitored over time.

- Final models are registered in a Model Registry and deployed via a model-serving layer for real-time integration with ERP services.

- Agentic workflows coordinate multi-step tasks and connect the ML pipeline with the conversational interface.

Combined System

- Together, Diagrams 3.1 and 3.2 represent an integrated ecosystem of microservices, ML processes, and agentic automation that delivers continuous intelligence across the ERP.
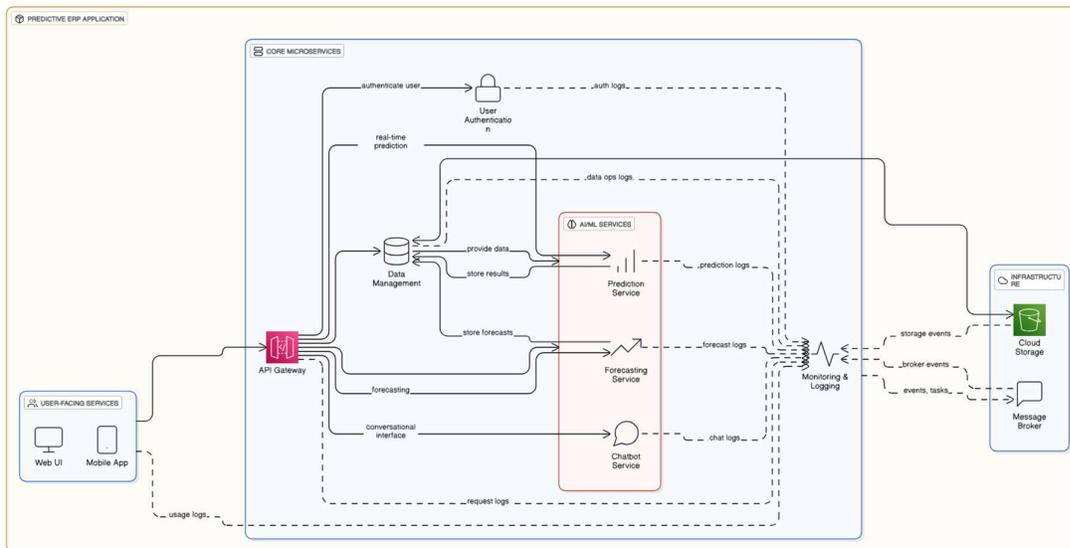

Figure 3.2 ERP architecture

4. Gap Finding & Future Scope

From the exploration of existing research conducted on the advancement of AI and ML systems that can be utilized in the development of a next generation ERP system, there are some considerable research gaps that need to be filled.

There is a need for standardized AI and ML integration in ERP workflows which makes the current systems rigid and unable to adapt quickly to changing business needs. And there is no uniformity and one workflow cannot be directly used in two different enterprises due to varying schemas and conventions.

The use of MLOps is still in its early stages and further developments in this workflow can improve the ease of maintaining production ML systems. There's a lack of studies and developments regarding the explainability and transparency of the reasoning and decisions made by AI agents. And also, research on real-time analytics needs to be matured in order to make a better architecture. Another research gap is about the shortage of high-quality datasets that we can use for training ERP specific ML models.

The future prospects of AI-powered next generation of ERP systems are very promising with exciting research advancements to fill these current gaps, which are currently ongoing in the industry. Enterprises and businesses have more than enough incentive to invest in research and development of modern ERP systems that can boost their productivity and maybe even revenue generation.

5. Conclusion

The addition of AI and ML models' capabilities in business operation by development of next generation ERP systems with AI capabilities will be a crucial step in advancing business growth in this modern age.

Our review of existing literature mentions the advancements in technology like MSA [4], MLOps [5, 6] and AI-powered agents and data pipelines [4, 9, 12] that can improve real-time analytics and predictions [3, 4, 7]. We also explored the research done towards the definition of transparency and explainability with eXplainable AI (XAI) feature metrics [13].

The proposed architecture discussed in Section III may be able to fill some of the research gaps discovered in this review and can potentially contribute to a standardized next generation ERP systems that provide value to enterprises and business in this AI era.

REFERENCES

[1] S. Kandhikonda, "AI-Powered ETL: Transforming Data with Smarter Pipelines," *Int. Res. J. Mod. Eng. Technol. Sci.*, vol. 7, no. 3, Mar. 2025.

[2] D. Kreuzberger *et al.*, "Machine Learning Operations (MLOps): Overview, Definition, and Architecture," *IEEE Access*, vol. 11, 2023.

[3] Sharma, V. K., Arnold, K. S., Mishra, P., & Sharma, H. (2021). Focaltheorem–A portfolio Web Application. Turkish Journal of Computer and Mathematics Education, 12(3), 3809-3814.

[4] I. Oumoussa and R. Saidi, "Evolution of Microservices Identification in Monolith Decomposition: A Systematic Review," *IEEE Access*, vol. 12, pp. 23389-23405, 2024.

[5] H. Yang *et al.*, "FinRobot: Generative Business Process AI Agents for Enterprise Resource Planning in Finance," *arXiv preprint* arXiv:2506.01423, Jun. 2025.

[6] E. Kandogan *et al.*, "Orchestrating Agents and Data for Enterprise: A Blueprint Architecture for Compound AI," *arXiv preprint* arXiv:2504.08148, Apr. 2025.

[7] W. Chen *et al.*, "Real-Time Analytics: Concepts, Architectures, and ML/AI Considerations," *IEEE Access*, vol. 11, pp. 71634-71657, 2023.

[8] A. Aljohani, "Predictive Analytics and Machine Learning for Real-Time Supply Chain Risk Mitigation and Agility," *Sustainability*, vol. 15, no. 20, art. 15088, Oct. 2023.

[9] A. B. Arrieta et al., "Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI" *Information Fusion*, vol. 58, pp. 82-115, Jun. 2020.

[10] Bandara, F., Jayawickrama, U., Subasinghage, M. *et al.*, "Enhancing ERP Responsiveness Through Big Data Technologies: An Empirical Investigation." *Inf Syst Front* 26, 251–275 (2024).

[11] P. Lech *et al.*, "Evolution and deployment options of Enterprise Systems – the case of SAP S/4HANA enterprise application suite," in *Artificial Intelligence and Machine Learning. IBIMA-AI 2024. Communications in Computer and Information Science*, vol. 2299, Nov. 2023.

[12] M. A. Mannan *et al.*, "Transforming ERP systems with collaborative AI: Paving the path to strategic growth and sustainability," *Array*, vol. 20, Art. no. 100517, Dec. 2025.

[13] M. Ganesh *et al.*, "An Analysis of the Significance of Spring Boot in The Market," in *Proc. 2022 Int. Conf. Inventive Comput. Technol. (ICICI)*, Jul. 2022.

[14] M. Golec *et al.*, "Music Emotion Recognition-Based Business-Oriented Visualization Framework Using AI-driven Serverless Cloud Computing," *Int. J. Bus. Anal.*, vol. 12, no. 1, May 2025.

[15] F. Brachten *et al.*, "The acceptance of chatbots in an enterprise context – A survey study," *Int. J. Inf. Manage.*, vol. 60, Art. no. 102375, Oct. 2021.

[16] V. Chang *et al.*, "Cloud Computing Adoption Framework–a security framework for business clouds," *Future Gener. Comput. Syst.*, vol. 57, Oct. 2015.

[17] H. S. Bhaskaran *et al.*, "Development of a cloud-based IoT system for livestock health monitoring using AWS and python," *Smart Agric. Technol.*, vol. 9, Art. no. 100524, Dec. 2024.

[18] J. Jiang *et al.*, "A deep learning framework integrating Transformer and LSTM architectures for pipeline corrosion rate forecasting," *Comput. Chem. Eng.*, vol. 204, Art. no. 109365, Jan. 2026.

[19] Hashmi, T., Chauhan, G., Kumar, N., Srivastava, A., & Sharma, H. (2024, December). Exploring Artificial Intelligence & Machine Learning in Precision Agriculture. In 2024 International Conference on Emerging Technologies and Innovation for Sustainability (EmergIN) (pp. 320-324). IEEE.

[20] W. De Michele *et al.*, "Automated Business Process Analysis: An LLM–Based Approach to Value Assessment," *arXiv preprint arXiv:2504.06600*, Apr. 2025.