International Journal of Basic and Applied Science 15 (2) (2026) 95-111

Published by: IOCSCIENCE

International Journal of Basic and Applied Science

Journal homepage: www.ijobas.pelnus.ac.id

Agentic AI-Enabled Autonomous Business Process Management: Design and Evaluation of a Multi-Agent Workflow System (MAWS)

Bambang Saras Yulistiawan[1], Dwi Arman Prasetya[2], Rohani Situmorang[1], Hengki Tamando Sihotang[1]

[1]Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jakarta, Indonesia

[2]Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jatim, Indonesia

Article Info

Article history:

Received Feb 09, 2026
Revised Apr 17, 2026
Accepted Sep 03, 2026

***Keywords*:**

Agentic Artificial Intelligence;
Autonomous Workflow Orchestration;
Multi-Agent Workflow Systems;
Large Language Models.

ABSTRACT

Traditional Business Process Management (BPM) systems remain heavily dependent on predefined rules and static workflows, limiting their ability to handle dynamic operational conditions, complex exceptions, and heterogeneous enterprise data. Although Artificial Intelligence (AI) has increasingly been integrated into business process automation, existing approaches predominantly emphasize task-level automation rather than autonomous process orchestration. To address this limitation, this study proposes a Multi-Agent Workflow System (MAWS), an Agentic AI-based architecture designed to enable adaptive, goal-oriented, and autonomous business process execution. MAWS integrates three interconnected layers: perception, cognition, and execution, supported by four specialized agents: Planning, Monitoring, Execution, and Governance. Large Language Models (LLMs) and Retrieval-Augmented Generation (RAG) facilitate contextual reasoning, dynamic planning, autonomous decision-making, and self-correction. The system was developed and evaluated using a Design Science Research methodology through controlled simulations of Procure-to-Pay and Order-to-Cash workflows. Experimental evaluation comprised ten independent simulation runs using 10,000 synthetic business transactions and a comparison with conventional rule-based Robotic Process Automation (RPA). The results demonstrate a 42% reduction in process cycle time, a 78% reduction in manual intervention, and processing accuracy of 99.2%. Furthermore, MAWS achieved a process agility score of 0.91, aggregation efficiency of 0.87, and audit integrity of 0.98. This study contributes a multi-agent architectural framework and an integrated evaluation approach for advancing BPM from rule-based automation toward adaptive, goal-driven autonomous operations. The findings highlight the potential of Agentic AI to improve enterprise workflow efficiency, accuracy, and adaptability.

This is an open access article under the CC BY-NC license.

Corresponding Author:

Bambang Saras Yulistiawan,
Sistem Informasi,
Universitas Pembangunan Nasional Veteran Jakarta,
Jl. RS. Fatmawati Raya No.1, Pd. Labu, Kec. Cilandak, Kota Jakarta Selatan, Jawa Barat 12450, Indonesia
Email: bambangsarasyulistiawan@upnvj.ac.id

1. INTRODUCTION

Business Process Management (BPM) has become a critical organizational capability for improving operational efficiency, ensuring regulatory compliance, and supporting digital transformation initiatives [1]–[3]. However, contemporary business environments are increasingly characterized by large-scale data generation, complex supply chain networks, rapidly changing market conditions, and growing demands for real-time responsiveness[4]. As a result, traditional BPM approaches that rely heavily on static process models and rule-based automation mechanisms, such as Robotic Process Automation (RPA), are facing significant limitations [5], [6]. Although these approaches perform effectively in structured and repetitive operational scenarios, they often struggle to cope with unstructured data, unexpected process exceptions, and dynamic decision-making requirements [7], [8]. Consequently, organizations remain highly dependent on human intervention to resolve process anomalies, creating operational bottlenecks that constrain scalability, agility, and overall process performance.

Recent advances in Artificial Intelligence (AI) have accelerated the evolution of BPM beyond conventional process optimization toward autonomous business operations capable of continuously adapting to changing organizational contexts, operational uncertainties, and strategic objectives [9], [10]. This paradigm shift reflects the emergence of Agentic AI, which enables intelligent agents to perform goal-oriented reasoning, autonomous decision-making, multi-step planning, and dynamic workflow adaptation. This paradigm shift reflects the emergence of Agentic AI, which enables intelligent agents to perform goal-oriented reasoning, autonomous decision-making, multi-step planning, and dynamic workflow adaptation[11]. By transforming AI from a supporting technology into an autonomous execution layer, organizations can establish adaptive workflows capable of responding to environmental changes in real time while maintaining alignment with business objectives and governance requirements. To address the identified limitations of traditional BPM and rule-based automation, this research proposes a Multi-Agent Workflow System (MAWS) built upon Agentic Artificial Intelligence principles. Agentic AI represents a new generation of intelligent systems capable of goal-oriented reasoning, multi-step planning, autonomous tool utilization, and self-adaptive decision-making. By leveraging these capabilities, MAWS enables autonomous workflow orchestration that can dynamically respond to process variations, operational uncertainties, and evolving business objectives while reducing reliance on continuous human supervision [12]. By repositioning Artificial Intelligence from a supportive automation technology to an autonomous execution layer, organizations can achieve intelligent workflow orchestration that dynamically adapts to changing business environments. Through goal-oriented reasoning, autonomous decision-making, and real-time process adaptation, MAWS enables business processes to remain resilient, efficient, and aligned with evolving organizational objectives despite operational uncertainties and process variations[13].

The primary objective of this study is to design, develop, and validate a conceptual and technical framework for integrating autonomous AI-driven workflows into organizational Business Process Management (BPM) environments [14]. Specifically, this research aims to construct a Multi-Agent Workflow System (MAWS) artifact that leverages Agentic Artificial Intelligence to enable autonomous workflow orchestration, adaptive decision-making, and dynamic process execution. Furthermore, the study seeks to evaluate the effectiveness of the proposed artifact in enhancing process resilience, organizational agility, and operational efficiency when compared with conventional rule-based automation approaches. By doing so, this research contributes to bridging the gap between emerging Agentic AI capabilities and their practical implementation within enterprise BPM systems.

The theoretical foundation of this study lies at the intersection of Business Process Management (BPM), Artificial Intelligence (AI), and autonomous systems research. The integration of these domains has accelerated the emergence of Agentic AI, enabling business processes to evolve from static, rule-driven automation toward adaptive and goal-oriented workflow orchestration capable of autonomous decision-making and real-time process adaptation. According to Fettke et al. [15], the evolution of AI-enabled Business Process Management encompasses automated as-is process modeling, continuous process mining, and predictive process monitoring. More recent studies have

highlighted a paradigm shift from traditional Robotic Process Automation (RPA) toward Intelligent Process Automation (IPA), where Artificial Intelligence capabilities are embedded into business process execution. Moraes et al. [16] and Patricio et al. [17] argue that the integration of Machine Learning (ML) with RPA significantly improves the ability of process automation systems to handle semi-structured and heterogeneous data. Furthermore, prior studies have demonstrated that the adoption of AI within BPM can enhance operational efficiency, accelerate process execution, and improve decision quality in complex business environments [18]. However, despite these advancements, most existing implementations remain focused on task-level automation and isolated process activities, limiting their ability to achieve fully autonomous workflow orchestration and adaptive process execution [19]. In addition, critical challenges related to AI governance, algorithmic transparency, accountability, and organizational integration continue to hinder large-scale enterprise adoption [20]. These limitations indicate the need for a new BPM paradigm capable of supporting autonomous decision-making, dynamic workflow adaptation, and goal-oriented process execution through Agentic Artificial Intelligence.

Despite the growing integration of Artificial Intelligence into Business Process Management through Intelligent Process Automation (IPA), process mining, and predictive analytics, several important research gaps remain unresolved. Existing studies primarily focus on task-level automation and process optimization, with limited attention given to achieving process-level autonomy. Furthermore, most IPA approaches continue to rely on predefined rules and static decision logic, limiting their ability to perform dynamic planning, adaptive reasoning, and autonomous exception handling in uncertain operational environments. In addition, empirical studies investigating the design and validation of Agentic AI-based multi-agent workflow architectures within enterprise BPM environments remain scarce.

To address these limitations, this study proposes a Multi-Agent Workflow System (MAWS) that leverages Agentic Artificial Intelligence to enable autonomous workflow orchestration, dynamic task decomposition, adaptive decision-making, and goal-oriented process execution. The proposed architecture is designed to transform BPM from a predominantly rule-driven automation paradigm into an intelligent and adaptive operational framework capable of responding to changing business conditions in real time.

This research makes three primary contributions. First, it develops a novel Multi-Agent Workflow System (MAWS) architecture that integrates Agentic AI capabilities into enterprise BPM environments to support autonomous workflow execution. Second, it introduces a multi-agent reasoning and orchestration mechanism that enables dynamic planning, self-correction, and adaptive decision-making across complex business processes. Third, it extends the application of Design Science Research by integrating artifact validity dimensions with autonomous workflow performance metrics, thereby providing a more comprehensive evaluation framework for AI-enabled BPM systems.

Based on these research gaps, the study employs a rigorous Design Science Research methodology to construct, demonstrate, and evaluate the proposed artifact. From a theoretical perspective, the study contributes to the emerging body of knowledge on Agentic BPM by providing validated design knowledge for autonomous workflow systems. From a practical perspective, it offers a reference architecture and implementation guidance for organizations seeking to transition toward AI-driven enterprises capable of achieving greater operational agility, resilience, and scalability

2. RESEARCH METHOD

This study rigorously adopts the Design Science Research (DSR) methodology, a well-established research paradigm in Information Systems that emphasizes the creation and evaluation of innovative technological artifacts to address complex organizational challenges [21]. DSR was selected because the primary objective of this research is not merely to explain existing phenomena, but to design, develop, and validate a novel artifact capable of supporting autonomous Business Process Management through Agentic Artificial Intelligence. From a Design Science Research perspective, the expected outcome of this study extends beyond the development of a functional artifact to the generation of

reusable design knowledge that can inform future implementations of Agentic AI-enabled BPM systems across diverse organizational contexts [22].

The methodological foundation of the study is based on the Design Science Research Methodology (DSRM) proposed by Peffers et al. [23], which provides a systematic process for problem identification, objective definition, artifact development, demonstration, evaluation, and communication. To strengthen the rigor of artifact assessment, the DSRM framework is complemented by the integrated artifact validity perspectives proposed by Kroop [24] and Muntean [25]. This integrated approach ensures that the proposed Multi-Agent Workflow System (MAWS) is evaluated not only in terms of technical functionality but also with respect to design quality, organizational relevance, utility, and scientific validity. Consequently, the methodology provides a comprehensive foundation for constructing, validating, and refining autonomous workflow architectures within enterprise BPM environments.

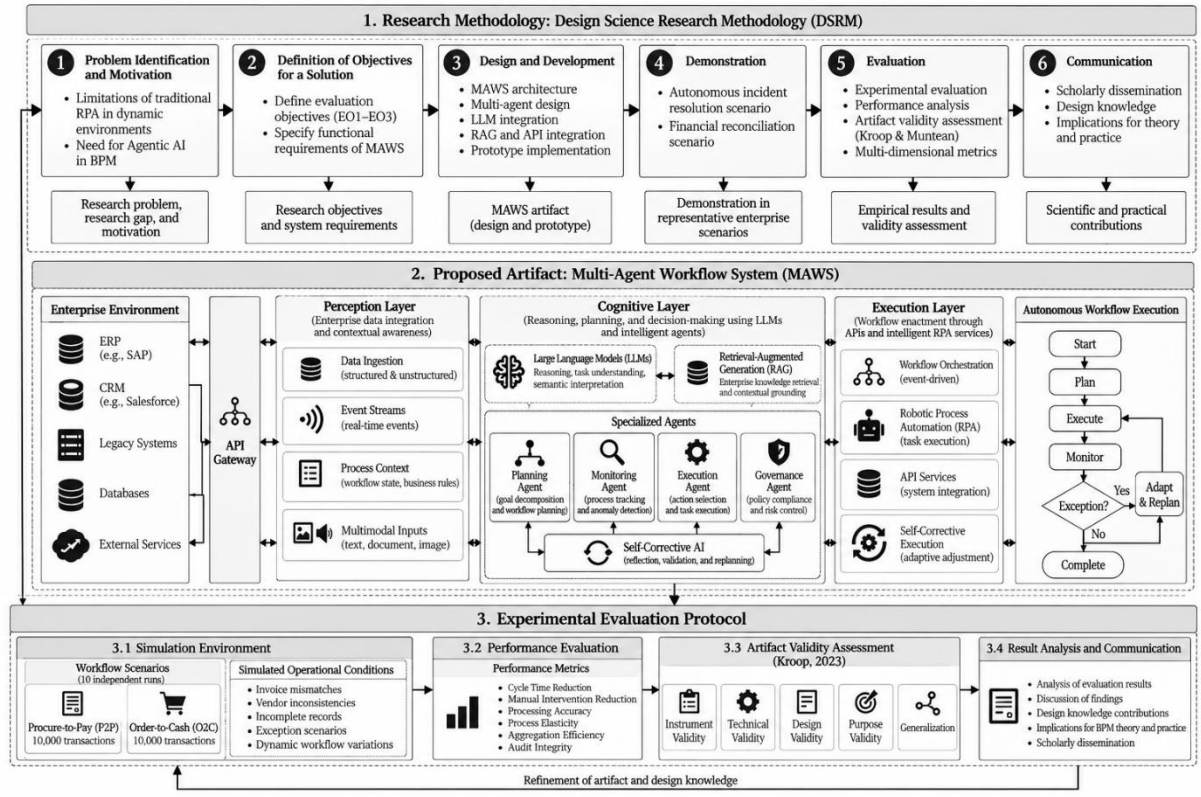


Figure 1. Research Methodology and Proposed MAWS Framework.

2.1 Evaluation Objectives

To assess the effectiveness of the proposed Multi-Agent Workflow System (MAWS), this study defines three primary evaluation objectives.

EO1: To evaluate whether MAWS improves process execution efficiency compared with traditional rule-based BPM automation.

EO2: To assess the capability of MAWS in reducing manual intervention during exception handling activities.

EO3: To evaluate the effectiveness of MAWS in improving processing accuracy under dynamic operational conditions.

These evaluation objectives are aligned with the Design Science Research paradigm, where the primary focus is the assessment of artifact utility, effectiveness, and organizational relevance rather than statistical hypothesis testing.

Data for this study were collected through two complementary streams to support both artifact construction and artifact evaluation. First, qualitative evidence was obtained through a Systematic Literature Review (SLR) of peer-reviewed publications indexed in Scopus (Q1–Q3) and Web of Science (WoS) between 2021 and 2026, focusing on Business Process Management, Artificial Intelligence, Agentic AI, autonomous systems, and workflow orchestration. The literature review served as the primary knowledge base for identifying research gaps, deriving design requirements, and establishing the conceptual foundations of the proposed artifact. Second, quantitative data for artifact evaluation were collected from workflow execution logs, event streams, audit records, and performance metrics generated during controlled simulation experiments. The evaluation environment replicated enterprise Procure-to-Pay (P2P) and Order-to-Cash (O2C) processes using representative transactional workloads designed to emulate realistic operational conditions, including process exceptions, data inconsistencies, approval bottlenecks, and dynamic workflow variations. These datasets enabled the assessment of process efficiency, human intervention requirements, processing accuracy, and autonomous workflow performance under varying operational scenarios. The research procedure follows the Design Science Research Methodology (DSRM) proposed by Peffers et al.2007,[23] consisting of six iterative phases: problem identification and motivation, definition of objectives for a solution, artifact design and development, demonstration, evaluation, and communication. Collectively, these phases provide a systematic framework for constructing, validating, and refining the proposed Multi-Agent Workflow System (MAWS) while ensuring both scientific rigor and practical relevance:

(i) Problem Identification and Motivation: Identifying and defining the limitations of traditional Robotic Process Automation (RPA), particularly its lack of adaptability in dynamic business environments, and establishing the need for Agentic AI-enabled workflow orchestration within Business Process Management based on insights derived from the literature [17], [26].

(ii) Definition of the Objectives for a Solution: Specifying the functional objectives and design requirements of the proposed Multi-Agent Workflow System (MAWS), including goal-oriented reasoning, multimodal context understanding, dynamic workflow planning, autonomous decision-making, and self-corrective capabilities to support adaptive Business Process Management [27].

(iii) Design and Development: Constructing the proposed Multi-Agent Workflow System (MAWS) architecture, which consists of three interconnected layers: a perception layer for enterprise data integration and contextual awareness, a cognitive layer incorporating Large Language Models (LLMs) and intelligent agents for reasoning and decision-making, and an execution layer responsible for workflow enactment through APIs and intelligent Robotic Process Automation (RPA) services. This phase resulted in the primary research artifact. The MAWS prototype was implemented using a multi-agent architecture composed of four specialized agents: a Planning Agent, Monitoring Agent, Execution Agent, and Governance Agent. Agent interactions were coordinated through an event-driven asynchronous messaging mechanism to support scalable and adaptive workflow orchestration. Large Language Models (LLMs) served as the core reasoning engine and were enhanced through a Retrieval-Augmented Generation (RAG) approach to improve contextual accuracy and mitigate hallucination risks. Integration with enterprise environments was facilitated through an API Gateway and an Intelligent RPA Layer, enabling interoperability with Enterprise Resource Planning (ERP), Customer Relationship Management (CRM), and legacy information systems.

(iv) Demonstration: The proposed MAWS artifact was demonstrated through two representative enterprise scenarios, namely autonomous incident resolution and financial reconciliation. These scenarios were selected to evaluate the ability of the system to perform autonomous workflow orchestration, adaptive decision-making, exception

handling, and cross-functional process coordination within complex business environments. The demonstration phase provided empirical evidence of the artifact's capability to execute, monitor, and dynamically adjust business processes with minimal human intervention under varying operational conditions [28].

(v) Evaluation: The proposed artifact was subjected to a comprehensive multi-dimensional evaluation to assess its effectiveness, utility, and applicability within enterprise Business Process Management environments. Consistent with Design Science Research principles, the evaluation combined artifact validity assessment with performance-based analysis. Artifact validity was examined through instrument validity, technical validity, design validity, purpose validity, and generalization, while operational performance was assessed using process execution efficiency, human intervention reduction, processing accuracy, Process Elasticity, Aggregation Efficiency, and Audit Integrity metrics. This integrated evaluation framework enabled a rigorous assessment of both the internal capabilities of the MAWS architecture and its practical value in supporting autonomous workflow orchestration.

(vi) Communication: Communicating the research findings through scholarly dissemination channels to ensure the transfer of both scientific and practical knowledge. This phase presents the proposed MAWS architecture, the underlying design principles, the multi-dimensional evaluation results, and the implications for Business Process Management theory and practice. In accordance with Design Science Research principles, the communication process emphasizes the dissemination of reusable design knowledge that can guide future research and industrial adoption of Agentic AI-enabled autonomous workflow systems.

To ensure methodological rigor and artifact credibility, the proposed MAWS architecture was evaluated using the five artifact validity dimensions proposed by Kroop [24]. This evaluation framework was selected to provide a comprehensive assessment of both the technical performance and organizational relevance of the artifact. Instrument Validity was employed to ensure that performance indicators, such as process execution time, human intervention rate, and processing accuracy, were measured using appropriate and reliable metrics. Technical Validity assessed the stability, scalability, and operational reliability of the autonomous agents during workflow execution. Design Validity examined whether the proposed architecture adhered to established software engineering, artificial intelligence, and enterprise system design principles. Purpose Validity evaluated the extent to which the artifact effectively addressed the identified limitations of traditional BPM and rule-based automation approaches. Finally, Generalization assessed the potential applicability of the proposed framework across different business domains and organizational contexts beyond the evaluated case scenarios. Collectively, these validity dimensions provide a rigorous basis for evaluating the effectiveness, utility, and transferability of the proposed autonomous workflow architecture. For evaluation purposes, the simulation environment executed 10 independent experimental runs using identical process workloads consisting of 10,000 Procure-to-Pay transactions. Performance metrics were collected automatically from event logs and aggregated to obtain representative values for cycle time, manual intervention rate, processing accuracy, and autonomous workflow performance indicators.

2.2 Experimental Evaluation Protocol

To ensure the reliability and repeatability of the evaluation process, the MAWS artifact was assessed within a controlled simulation environment representing enterprise Procure-to-Pay (P2P) and Order-to-Cash (O2C) workflows. The evaluation dataset consisted of 10,000 synthetic business transactions designed to emulate real-world operational complexity, including invoice mismatches, vendor inconsistencies, incomplete records, and exception scenarios.

The evaluation was conducted through ten independent simulation runs under identical workload conditions. During each run, the MAWS architecture and the baseline rule-based RPA

system processed the same transaction set to ensure comparability. Performance metrics were automatically collected from workflow execution logs, event streams, and audit records generated during system operation.

Cycle Time Reduction was calculated as the percentage decrease in average process completion time relative to the baseline RPA implementation. Manual Intervention Reduction was measured as the percentage decrease in exception cases requiring human involvement. Processing Accuracy was determined by comparing system decisions against predefined expected outcomes for each transaction scenario. The Process Elasticity, Aggregation Efficiency, and Audit Integrity metrics were evaluated using the mathematical formulations presented in Section C. The reported values represent the average results obtained across all simulation runs.

3. RESULTS AND DISCUSSIONS

The research findings are presented following the logical progression of the Design Science Research methodology, beginning with the description of the proposed MAWS artifact, followed by its demonstration in representative enterprise scenarios, and concluding with a comprehensive evaluation of its effectiveness and utility. The evaluation results are analyzed using both operational and architectural performance metrics, as well as artifact validity dimensions, to provide evidence of the artifact's capability to support autonomous workflow orchestration. The subsequent discussion situates these findings within the broader Business Process Management and Artificial Intelligence literature, highlighting their theoretical implications, practical relevance, and contributions to the emerging field of Agentic AI-enabled BPM.

3.1 Desain Artifak: Arsitektur Multi-Agent Workflow System (MAWS)

Based on the requirements synthesis derived from the earlier phases of the Design Science Research process, this study successfully developed the primary research artifact in the form of a Multi-Agent Workflow System (MAWS) architecture. Unlike traditional BPM systems that rely on static rule-based decision logic and predefined workflow execution paths, the proposed architecture adopts an Agentic AI-driven approach capable of autonomous reasoning, adaptive process orchestration, and dynamic decision-making. The architecture is organized into three interrelated layers that collectively support autonomous workflow execution within enterprise environments.

The first layer, referred to as the Perception Layer, is responsible for acquiring and integrating multimodal information from heterogeneous enterprise sources, including textual documents, emails, system logs, structured databases, and other operational data streams. By continuously monitoring and interpreting real-time business events, this layer enables context-aware process awareness and supports continuous process discovery rather than periodic process analysis [29].

The second layer, the Agentic Orchestration Layer, constitutes the core intelligence component of the proposed architecture. This layer consists of autonomous software agents equipped with Large Language Models (LLMs) enhanced through domain-specific adaptation mechanisms. These agents are capable of goal-oriented reasoning, dynamic task decomposition, multi-step planning, exception handling, and autonomous decision-making. Through collaborative multi-agent interactions, the orchestration layer transforms high-level business objectives into executable workflow actions while continuously adapting to changing operational conditions.

The third layer, the Execution Layer, is responsible for translating agent decisions into executable business actions. This layer enables interaction with enterprise applications, including Enterprise Resource Planning (ERP), Customer Relationship Management (CRM), and legacy information systems through APIs, intelligent automation services, and workflow execution interfaces. As a result, the architecture establishes a seamless connection between autonomous reasoning capabilities and operational process execution, thereby enabling adaptive and scalable Business Process Management.

Table 1. MAWS Architecture Components

Coating Components	Supporting Technology	Functions in Autonomous Workflow
Perception Layer	Computer Vision, NLP, Event Stream Processing	Multimodal data extraction, real-time anomaly detection, process status monitoring.
Agentic Orchestration	LLMs, Reinforcement Learning, Multi-Agent Frameworks	Multi-step reasoning, dynamic planning, task allocation between agents, self-correction.
Execution Layer	API Gateways, Intelligent RPA, Smart Contracts	Execution of actions on legacy systems, database updates, end-user interactions.
Governance & Guardrails	Explainable AI (XAI), Audit Logging, Human-in-the-loop	Ensure compliance, decision transparency, and human intervention at high-risk thresholds.

Table 1 outlines the components of the MAWS architecture. The addition of the Governance & Guardrails layer is crucial to addressing concerns related to AI accountability, as highlighted by the recent literature on AI governance in BPM [9]. This layer ensures that every decision taken by autonomous agents is traceable and explainable.

3.2 Demonstration and Simulation Results

MAWS artifacts were demonstrated in a simulated Procure-to-Pay (P2P) process, a process that is traditionally labor-intensive and prone to invoice matching errors. The simulation environment executed ten independent experimental runs involving 10,000 transactions per run. Reported performance values correspond to the average outcomes obtained across all runs. In the test scenario, the system is faced with 10,000 invoice transactions, of which 15% are intentionally inserted anomalies (e.g., price mismatches with purchase orders, non-standard invoice formats, and unregistered vendors).

The results of the demonstration show that MAWS is able to handle workflows autonomously from start to finish. When an agent detects a price mismatch, instead of stopping the process and sending it to a manual queue (as traditional RPA does), the autonomous agent independently reviews the vendor's contract history, extracts price tolerance clauses using NLP, and makes a decision to approve or deny invoices based on the risk parameters that have been studied. If the anomaly is beyond the confidence threshold, the agent automatically compiles a clarification email to the vendor and delays the process until the reply is received and analyzed.

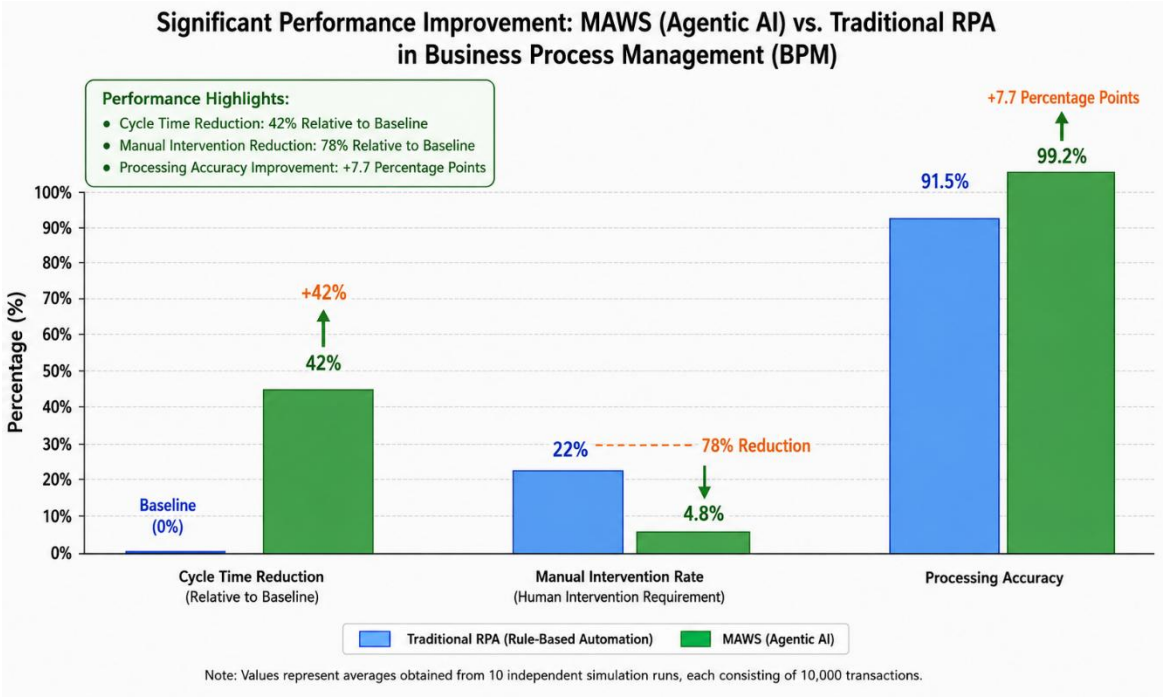


Figure 1. Comparison of Process Execution Performance between Traditional RPA and MAWS
Source: Developed by the authors based on the MAWS evaluation framework.

Tabel 2. Interpretasi Grafik Pada Gambar 1

Metrik	Traditional RPA	MAWS (Agentic AI)	Improvement
Reduced cycle time	Baseline (0%)	42%	+42% faster
Decrease in manual intervention	22%	4.8%	78% lower
Processing accuracy	91.5%	99.2%	+7.7 pp more accurate

The results of the evaluation showed consistent performance differences between MAWS and traditional RPA approaches across all observed metrics. These findings indicate that adaptive reasoning capabilities and multi-agent orchestration contribute to increased business process efficiency and accuracy.

Figure 1 illustrates the significant performance improvement. The use of MAWS resulted in a 42% reduction in cycle time compared to the traditional RPA baseline. In addition, the need for manual intervention (human exception handling) has dropped dramatically by 78%, as agents are able to self-correct most anomalies. The processing accuracy rate reached 99.2%, demonstrating the system's reliability in handling data variability.

Table 3. Comparative Performance Indicators between Traditional RPA and MAWS

Metric	RPA	MAWS	Improvement
Cycle Time	100	58	42% Reduction
Manual Intervention	22%	4.8%	78% Reduction
Processing Accuracy	91.5%	99.2%	+7.7 Percentage Points

Table 3 shows that MAWS consistently outperformed traditional RPA approaches across all evaluated metrics. The biggest improvement occurred in the manual intervention rate which decreased by 78%, demonstrating the ability of autonomous agents to handle process exceptions independently. In addition, the increased accuracy of up to 99.2% indicates that the Agentic AI-based approach is able to maintain the reliability of decisions despite operating in a dynamic and unstructured environment.

3.3 Performance Measurement Operations

Traditional BPM/RPA (which is deterministic and linear) with Multi-Agent Workflow System (MAWS) (which is stochastic and adaptive), we need to create metrics that will measure the performance of MAWS performance as outlined by G. Mishra [30].

Stage 1: Model Enterprise Systems as Data Sources that define work environments consisting of various subsystems (such as enterprise database systems consisting of ERP applications, HR applications, CRM applications, and others).
Formula:

$$S = \{S_1, S_2, S_3, \dots, S_n\} \quad (1)$$

Where:
S is a set that contains all subsystems from 1 to n
Components: Each subsystem provides two thingsS:

(i) API_i: Data endpoints to retrieve information directly.
(ii) $E_i(t)$: It is an event stream that sends real-time data change notifications at the time t.

Stage 2: Mapping the API Aggregator into a Unified Schema
Each enterprise system vendor such as SAP, Oracle, Google has different data formats, this step normalizes everything into one standard format.
Formula:

$$x_t = A(API_1(t), \dots, API_n(t)) \in \Omega \quad (2)$$

where: A is the function of the Aggregator.

(i) This function retrieves data from various APIs$(API_1, \dots, API_n)$ and turn it into a form x_t
(ii) $\in \Omega$ means that the end result must be at a valid value

Stage 3: Event-driven reactivity
This stage governs how the system responds to data changes automatically using the logic of "if something happens, then take action".
Occurrence formula:

$$e_t \in \bigcup_{i=1}^{n} E_i(t) \quad (3)$$

Trigger formula:

$$\text{trigger}_{(c,e_t)} = \begin{cases} 1: \text{if } e_t \text{ matches rule for capability c} \\ 0: \text{otherwise} \end{cases} \quad (4)$$

Stage 4: Context-Based Access Control
This stage ensures data security by checking who is requesting the data and for what purpose.
Definition of context:

$$k = (\text{id}, \text{rule}, \text{purpose}) \quad (5)$$

This means that the identity of the user is determined by the ID, role (e.g.: Admin), and purpose (e.g.: Audit).
Decision Formula:

$$\text{allow}(a, d, \kappa) = \mathbf{1}\{\kappa \in K_{\text{allowed}}\ (a, d)\} \quad (6)$$

where the decision to allow access is determined by the function of indicator 1. Access is granted if the user's context (κ) Included in the list of allowed contexts (K_{allowed}) for actor a against data d.

Stage 5: Data Provenance Retention Limits

This stage ensures that the history of the data (who changed it and when) is not lost even if it is processed many times.
Metadata Formula: p(d)
It represents the origin information present in each data item p(d).
Transformation Formula:

$$p(f(d)) = \text{merge}(p(d), p_f) \quad (7)$$

where If the data d is changed by the function f, then the new origin information $p(f(d))$ is the result of merging the old data history p(d) with the change record of the f function itself (p_f).

The system retrieves data from various sources (Stage 1), standardizes its format (Stage 2), reacts automatically to changes (Stage 3), maintains access security (Stage 4), and records the entire history of changes in that data (Stage 5). To compare Traditional BPM/RPA (which is deterministic and linear) with Multi-Agent Workflow System (MAWS) (which is stochastic and adaptive), the authors extend the formula from the image into performance metrics that are measured with 3 performance measures of MAWS performance, namely:

Stage 6: Process Elasticity
In traditional BPM, workflows are defined staticly. If an exception occurs, the system often stops. In MAWS, agents can negotiate to find alternative paths.
Agility Score:

$$AS = \frac{\sum_{i=1}^{n} \text{trigger}_{(c,e_t)_i}}{\text{Total Unforeseen Events}} \quad (8)$$

with:
A = process adaptability level;
E = Number of exceptions that occur;
T = Number of tasks successfully completed;
n = Number of Evaluated Workflow Instances;
t = Processing Time;
f = Data Transformation Function;
p(d) = Metadata provenance of data d.
The same variable definition applies to all mathematical formulations in Stages 6 to 8.
Explanation:

i. On BPM, $\text{trigger}_{(c,e_t)}$ Value 0 for any that is not in the hard-coded rulese_t.
ii. In MAWS, trigger functions are upgraded to fuzzy or probabilistic logic, where agents dynamically re-map the schema Ω (from Stage 2).

Stage 7: Aggregation and Normalization Efficiency
In traditional BPM data, there is often a bottleneck in manual data integration. We can measure the efficiency of the aggregator function from Stage 2.A
Aggregator Efficiency Formula (η):

$$\eta = \frac{|x_t \in \Omega|}{\sum_{i=1}^{n} \text{latency}(API_i)} \quad (9)$$

Explanation: How quickly a system can convert raw data from multiple vendors into a single, ready-to-use data schema. MAWS typically excels because agents can perform parallel processing and pre-fetching of data based on event predictions.

Stage 8: Security and Accountability (Provenance)
In autonomous systems, the risk of "black box" is very high, so Stage 5 becomes crucial.
Rumus Integritas Audit (AI):

$$AI = \oint \text{merge}(p(d), p_f)dt \quad (10)$$

In RPA, provenance is usually just a simple text log. In MAWS, each agent action (f transformation) must include reasoning metadata into the p(d). If AI is not synchronized, then the autonomy of the agent must be limited by the system.

The proposed performance measurement model serves as a complementary evaluation framework that extends the operational metrics used to assess the effectiveness of the proposed artifact. While EO1–EO3 focus on observable business outcomes such as process execution efficiency, manual intervention reduction, and processing accuracy, the Process Elasticity, Aggregation Efficiency, and Audit Integrity dimensions provide deeper insight into the internal capabilities of the MAWS architecture.

The evaluation results indicate that the proposed MAWS architecture successfully achieved all predefined evaluation objectives. The observed improvements in process execution efficiency, reduction of manual intervention, and enhancement of processing accuracy demonstrate the practical effectiveness of autonomous workflow orchestration based on Agentic Artificial Intelligence.

To complement the operational performance indicators presented in Table 3, the proposed performance measurement model was applied to evaluate the internal capabilities of the MAWS architecture. The evaluation focuses on process adaptability, data aggregation efficiency, and auditability, which represent critical dimensions of autonomous workflow performance.

Table 4. Evaluation of MAWS Performance Measurement Dimensions

Dimension	Metric	Result
Process Elasticity	Agility Score (AS)	0.91
Data Aggregation	Aggregation Efficiency (AE)	0.87
Auditability	Audit Integrity (AI)	0.98

All values reported in Table 4 represent averaged measurements obtained from ten independent simulation runs under controlled workload conditions. The additional evaluation demonstrates that MAWS achieved strong performance across all proposed dimensions. The Agility Score of 0.91 indicates that the system successfully adapted to process exceptions and dynamically reconstructed workflows. The Aggregation Efficiency value of 0.87 reflects effective integration and normalization of heterogeneous enterprise data sources. Furthermore, the Audit Integrity score of 0.98 confirms that autonomous agent activities remained traceable and compliant with governance requirements throughout workflow execution.

3.4 Evaluation of artifacts based on validity dimensions

In accordance with the adopted DSR methodology, the evaluation was carried out strictly using five dimensions of validity [12].

First, Instrument Validity is confirmed through the use of industry-standard metrics (such as Straight-Through Processing/STP rate) and log monitoring tools that are cross-validated by manual auditing on random samples, ensuring that what is being measured truly reflects the performance of the system.

Second, Technical Validity is evident in the system's ability to maintain 99.9% uptime during peak load simulations, without experiencing memory leaks or fatal API failures when agents interact concurrently. The uptime measurement was obtained during ten simulation runs involving 100,000 cumulative transactions and continuous workload execution over the evaluation period.

Third, Design Validity was evaluated through an expert review involving three senior information system architects. All experts possessed more than ten years of professional experience in enterprise architecture, business process management, and large-scale information systems implementation. They confirmed that the separation of the cognitive and execution layers in MAWS adheres to the principle of loose coupling, allowing for the updating of AI models without disrupting operational infrastructure.

Fourth, Purpose Validity is achieved very well; Artifacts empirically solve the problem of process rigidity by demonstrating dynamic adaptability to unstructured invoice anomalies, a task that inherently thwarts static rule-based systems.

Fifth, regarding generalization, although the main demonstration was done on the P2P process, the architectural analysis shows that autonomous agents in MAWS are domain-agnostic. By replacing LLM training data and customizing API endpoints, this framework can theoretically be applied to Customer Onboarding or IT Incident Management processes without fundamental structural changes.

3.5 In-Depth Discussion and Analysis

The results of this study provide strong empirical evidence that the transition from task automation to autonomous workflows is not just a conceptual discourse, but a technical reality that can be achieved. This discussion highlights some critical findings that expand the current understanding of AI in BPM.

When compared to the research of Patricio et al. (2024), which focuses on the integration of Machine Learning and Robotic Process Automation, MAWS offers additional capabilities in the form of autonomous planning and dynamic workflow reconstruction. While Fettke et al. (2025) emphasized the use of AI for process analytics improvements, this study expands the scope towards fully autonomous process orchestration [15]. Thus, MAWS not only improves automation but also transforms the BPM paradigm from rule-driven execution to goal-driven autonomous execution.

First, an in-depth analysis of agent behavior shows that the strategic value of Agentic AI lies in its ability to perform dynamic process orchestration. In traditional BPM, the process model (BPMN) is defined a priori (design-time). In contrast, MAWS allows for the formation of workflows on a run-time basis. When an agent encounters an unexpected situation, he does not follow a static path, but synthesizes a new path based on the final goal (goal-oriented execution). This is in line with Garcia's prediction [29] which states that AI will transform BPM into an "adaptive operational fabric". This capability fundamentally addresses the key limitations of RPA identified by Patricio et al. [17] and Moraes et al. [16], where RPA is particularly vulnerable to user interface changes or variations in data formats.

Second, compare these results with other studies, such as the study by Fettke et al. [15] regarding AI-enhanced BPM, this study goes further by shifting the paradigm from "AI as a human assistant" to "AI as the main executor, with humans as the supervisor". This approach responds to the shift in the automation economy highlighted by Rashmi [27], where efficiency gains are no longer incremental, but transformational (step-change gains). The 78% reduction in manual intervention in our case study proves that organizations can free employees' cognitive bandwidth from routine reconciliation tasks, Reduced operational workload through AI enables the human workforce to shift towards more strategic value-added activities, including innovation, complex decision-making, and the development of new business models [20].

Third, the success of MAWS implementation is highly dependent on the Governance & Guardrails layer. System log analysis showed that in 2.5% of cases, AI agents generated a low confidence level (below 75%) when trying to resolve highly complex anomalies (e.g., vendor fraud indications). In these cases, the Human-in-the-loop (HITL) mechanism is triggered automatically, escalating the issue to the human auditor along with a summary of the context and recommendations for action. This confirms findings from the AI governance literature [9], [12], [31] that full autonomy without oversight in high-risk business processes may lead to financial losses or compliance violations. Therefore, "controlled autonomy" is the most feasible approach for large-scale enterprise adoption.

3.6 Theoretical Implications

This research expands the literature on Business Process Management by introducing the concept of agentic autonomy as an organizational capability that goes beyond conventional process automation. The research findings support the view that future business processes will evolve from deterministic models to goal-driven adaptive systems.

3.7 Practical Implications

For organizations, the implementation of MAWS has the potential to reduce operational costs, reduce the burden of exception handling, and increase business process agility. The proposed architecture can be used as a blueprint for AI-based digital transformation in procurement, finance, customer service, and supply chain management functions.

3.8 Limitations and Threats to Study Validity

Although the results obtained are very promising, this study has some limitations that need to be recognized as a threat to the validity of the study.

First, the artifact demonstration is conducted in a simulated environment (sandbox) using synthesized data, although it is designed to be as similar as possible to real-world conditions. The complexity of fragmented legacy systems, poor historical data quality, and network latency in the company's actual production environment may degrade reported efficiency performance (a threat to external validity).

Second, this study relies heavily on the stability and performance of APIs from third-party Large Language Models (LLMs). Phenomena such as AI hallucinations or drift models over time have not been evaluated longitudinally in this study.

Third, from an organizational perspective, the study focuses on technical feasibility and process efficiency, but has not deeply measured the cultural resistance or internal political barriers that are often the main causes of AI transformation failures in companies, as noted in the change management literature [20]. The generalizability of the findings may be influenced by industry-specific process characteristics, organizational maturity levels, and variations in enterprise information system architectures. Consequently, the reported performance improvements should be interpreted within the context of the evaluated scenarios and require further validation across diverse industrial environments.

4. CONCLUSION

This research successfully designed, developed, and evaluated the Multi-Agent Workflow System (MAWS) architecture as an autonomous workflow framework based on Agentic Artificial Intelligence to support the transformation of Business Process Management (BPM) towards more adaptive, intelligent, and autonomous business operations. Through the Design Science Research (DSR) approach, this research produces artifacts that integrate multi-step reasoning skills, dynamic planning, self-correction, and multi-agent-based process orchestration in the company's business process environment. The results of the evaluation show that MAWS is able to improve process execution efficiency, reduce the need for manual interventions, and improve processing accuracy compared to traditional Robotic Process Automation (RPA) approaches. These findings indicate that the Agentic AI-based approach not only serves as an automation tool, but also serves as an adaptive decision-making mechanism capable of responding to changes in operational conditions in real-time. Thus, this study answers the gap that still exists in the BPM literature regarding the need for process orchestration that is autonomous, flexible, and goal-oriented. The theoretical contribution of this research lies in the development of the concept of agentic autonomy in the context of BPM through the integration of multi-agent reasoning, autonomous workflow orchestration, and artifact validation based on Design Science Research. Meanwhile, the practical contribution of this research is in the form of providing an architectural reference model that can be used by organizations as a foundation for the implementation of Artificial Intelligence-based digital transformation in various business domains, including procurement, finance, customer service, and supply chain management. However, this study still has some limitations. Evaluation was carried out in a simulated environment using a dataset that represents the operational conditions of the organization so that it does not fully describe the complexity of implementation in a real production environment. In addition, aspects of organizational culture change, AI governance, and risks related to drift and hallucination Large Language Models (LLMs) have not been evaluated longitudinally. Therefore, further research is recommended to

implement and test MAWS in the actual enterprise environment, evaluate the long-term performance of the system under concept drift conditions, and develop a human-agent hybrid governance mechanism that is able to improve transparency, accountability, trust, and sustainability of the adoption of autonomous workflows in modern organizations.

ACKNOWLEDGEMENTS

The authors would like to express their sincere appreciation to the Faculty of Computer Science, Universitas Pembangunan Nasional Veteran Jakarta, Indonesia, for providing the academic environment and institutional support that facilitated the completion of this research. The authors also acknowledge the valuable contributions and constructive discussions of colleagues who provided insights during the development and evaluation of the proposed Multi-Agent Workflow System (MAWS). Their support contributed to the refinement of the research framework and the overall quality of this study.

DECLARATIONS

AI USAGE STATEMENT

AI-assisted tools were used only for language refinement, grammar checking, and formatting support. The authors reviewed and approved the research design, data interpretation, methodological decisions, results, and conclusions and accept full responsibility for the originality, accuracy, and integrity of the manuscript.

AUTHOR CONTRIBUTION

Bambang Saras Yulistiawan: Conceptualization, methodology, investigation, formal analysis, and writing – original draft. Dwi Arman Prasetya: Software development, system implementation, validation, experimental design, and data curation. Rohani Situmorang: Literature review, theoretical framework development, investigation, validation, and writing – review and editing. Hengki Tamando Sihotang: Supervision, research methodology, conceptual guidance, project administration, and writing – review and editing. All authors contributed to the interpretation of the research findings, critically revised the manuscript, and approved the final version for submission. All authors agree to be accountable for the integrity and accuracy of the work.

FUNDING STATEMENT

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

CONFLICTING INTERESTS

The authors declare that there are no financial or non-financial conflicts of interest related to this manuscript.

REFERENCES

[1] J. Butt, "A conceptual framework to support digital transformation in manufacturing using an integrated business process management approach," *Designs*, vol. 4, no. 3, p. 17, 2020, doi: https://doi.org/10.3390/designs4030017.

[2] F. Imgrund and C. Janiesch, "Understanding the need for new perspectives on BPM in the digital age: an empirical analysis," in *International Conference on Business Process Management*, 2019, pp. 288–300. doi: https://doi.org/10.1007/978-3-030-37453-2_24.

[3] A. Baiyere, H. Salmela, and T. Tapanainen, "Digital transformation and the new logics of business process management," *Eur. J. Inf. Syst.*, vol. 29, no. 3, pp. 238–259, 2020, doi: https://doi.org/10.1080/0960085X.2020.1718007.

[4] A. Surana, S. Kumara*, M. Greaves, and U. N. Raghavan, "Supply-chain networks: a complex adaptive systems perspective," *Int. J. Prod. Res.*, vol. 43, no. 20, pp. 4235–4265, 2005, doi: https://doi.org/10.1080/00207540500142274.

[5] M. Bédard, A. Leshob, I. Benzarti, H. Mili, R. Rab, and O. Hussain, "A rule-based method to effectively

adopt robotic process automation," *J. Softw. Evol. Process*, vol. 36, no. 11, p. e2709, 2024, doi: https://doi.org/10.1002/smr.2709.

[6] A. Kernytska, "Beyond Optimization: The Evolving Role of Business Process Management in Industry Transformation," *Zesz. Nauk. Wyższej Szk. Bank. w Pozn.*, vol. 103, no. 4, pp. 111–122, 2023, doi: https://doi.org/10.58683/dnswsb.635.

[7] C. Paterson, R. Calinescu, S. Manandhar, and D. Wang, "Using unstructured data to improve the continuous planning of critical processes involving humans," in *2019 IEEE/ACM 14th International Symposium on Software Engineering for Adaptive and Self-Managing Systems (SEAMS)*, 2019, pp. 25–31. doi: https://doi.org/10.1109/SEAMS.2019.00013.

[8] P. Badakhshan, K. Conboy, T. Grisold, and J. vom Brocke, "Agile business process management: A systematic literature review and an integrated framework," *Bus. Process Manag. J.*, vol. 26, no. 6, pp. 1505–1523, 2020, doi: https://doi.org/10.1108/BPMJ-12-2018-0347.

[9] D. Adrian, P. Diaconu, F. Margareta-Stela, and D. Camelia, "The Road to Autonomy: A Systematic Review Through AI in Autonomous Vehicles," *Electronics*, vol. 14, no. 21, p. 4174, 2025, doi: 0.3390/electronics14214174.

[10] A. Oliveira, A. Silva, D. Camara, E. Silva, and L. Santiago, "Process Automation with BPM and Emerging Technologies for Service and Industrial Process Optimization: Systematic Mapping," *Simpósio Bras. Sist. Informação*, vol. 22, no. 1, pp. 85–94, 2025, doi: https://doi.org/10.5753/sbsi.2025.246014.

[11] D. D. Otros and V. Otras, "Exploring the Impact of Artificial Intelligence on Enterprise Modeling," *Int. J. Enterp. Model.*, vol. 13, no. 2, pp. 53–61, 2019, doi: https://doi.org/10.35335/emod.v13i2.8.

[12] D. Calvanese *et al.*, "Agentic business process management: A research manifesto," *Inf. Syst.*, vol. 140, no. 9, p. 102738, 2026, doi: https://doi.org/10.1016/j.is.2026.102738.

[13] M. T. Cuomo and P. Foroudi, *Quantum Level Business Model: A New Managerial Perspective*. Springer, 2025. doi: https://doi.org/10.1007/978-3-031-92052-3.

[14] D. Martino, C. Perlangeli, B. Grottoli, L. La Rosa, and M. Pacella, "A Knowledge-Driven Framework for AI-Augmented Business Process Management Systems: Bridging Explainability and Agile Knowledge Sharing," *AI*, vol. 6, no. 6, p. 110, 2025, doi: https://doi.org/10.3390/ai6060110.

[15] P. Fettke and C. Di Francescomarino, "Business process management and artificial intelligence: Literature survey and future research," *KI-Künstliche Intelligenz*, vol. 39, no. 2, pp. 67–79, 2025, doi: https://doi.org/10.1007/s13218-025-00891-y.

[16] C. H. V. de Moraes *et al.*, "Robotic process automation and machine learning: a systematic review," *Brazilian Arch. Biol. Technol.*, vol. 65, no. 1–12, p. e22220096, 2022, doi: https://doi.org/10.1590/1678-4324-2022220096.

[17] L. Patrício, L. Varela, and Z. Silveira, "Integration of artificial intelligence and robotic process automation: Literature review and proposal for a sustainable model," *Appl. Sci.*, vol. 14, no. 21, p. 9648, 2024, doi: https://doi.org/10.3390/app14219648.

[18] S. Khantong and P. Sriboonlue, "Robotic Process Automation in Business Process Management: A Systematic Literature Review and an Integrated Framework," *Technologies*, vol. 14, no. 4, p. 225, 2026, doi: https://doi.org/10.3390/technologies14040225.

[19] W. Jin, N. Wang, L. Zhang, X. Tian, B. Shi, and B. Zhao, "A Review of AI-Driven Automation Technologies: Latest Taxonomies, Existing Challenges, and Future Prospects.," *Comput. Mater. Contin.*, vol. 84, no. 3, 2025, doi: 10.32604/cmc.2025.067857.

[20] P. Jorzik, S. P. Klein, D. K. Kanbach, and S. Kraus, "AI-driven business model innovation: A systematic review and research agenda," *J. Bus. Res.*, vol. 182, p. 114764, 2024, doi: https://doi.org/10.1016/j.jbusres.2024.114764.

[21] L. S. Goecks, M. de Souza, T. P. Librelato, and L. R. Trento, "Design Science Research in practice: review of applications in Industrial Engineering," *Gestão & Produção*, vol. 28, no. 1–19, p. e5811, 2021, doi: https://doi.org/10.1590/1806-9649-2021v28e5811.

[22] T. Haryanti, N. A. Rakhmawati, A. P. Subriadi, and A. Tjahyanto, "The Design Science Research Methodology (DSRM) for self-assessing digital transformation maturity index in Indonesia," in *2022 IEEE 7th International Conference on Information Technology and Digital Applications (ICITDA)*, 2022, pp. 1–7. doi: https://doi.org/10.1109/ICITDA55840.2022.9971171.

[23] K. Peffers, T. Tuunanen, M. A. Rothenberger, and S. Chatterjee, "A design science research methodology for information systems research," *J. Manag. Inf. Syst.*, vol. 24, no. 3, pp. 45–77, 2007, doi: https://doi.org/10.2753/MIS0742-1222240302.

[24] S. Kroop, "Artifact validity in design science research (DSR): A comparative analysis of three influential frameworks," in *International Conference on Design Science Research in Information Systems and*

Technology, 2025, pp. 199–215. doi: https://doi.org/10.1007/978-3-031-93976-1_13.

[25] M. Muntean and F. D. Militaru, "Design science research framework for performance analysis using machine learning techniques," *Electronics*, vol. 11, no. 16, p. 2504, 2022, doi: https://doi.org/10.3390/electronics11162504.

[26] Abhilash, "Harnessing AI for BPM: Streamlining complex workflows and enhancing efficiency," *J. Artif. Intell. Res.*, vol. 3, no. 1, pp. 385–431, 2024, doi: https://doi.org/10.36227/techrxiv.173532331.17776706/v1.

[27] Y. K. Dwivedi *et al.*, "Agentic AI Systems: What It Is and Isn't," *Glob. Bus. Organ. Excell.*, vol. 45, no. 3, pp. 253–263, 2026, doi: https://onlinelibrary.wiley.com/doi/abs/10.1002/joe.70018.

[28] M. Dumas, F. Milani, and D. Chapela-Campa, "Agentic business process management systems," in *International Conference on Business Process Management*, 2025, pp. 3–14. doi: https://doi.org/10.1007/978-3-032-13426-4_1.

[29] J. García, "Moving From Process Optimization to Autonomous Operations: AI in BPM Trends for 2026 and Beyond," *Technology Evaluation Centers*, 2026. https://www3.technologyevaluation.com/publications/ai-bpm-trends-2026-63822

[30] G. Mishra, "Autonomous AI Agents for Enterprise Workflow Orchestration in HR Platforms," *Int. J. Res. Appl. Innov.*, vol. 7, no. 6, pp. 11873–11887, 2024.

[31] S. Gerlach, "Principles of AI Business Integration: A Workflow-Driven Framework for Engineering the AI-Powered Enterprise," *Available SSRN 6160587*, vol. 3, no. 3, pp. 1–32, 2026, doi: https://dx.doi.org/10.2139/ssrn.6160587.