Int. j. adv. multidisc. res. stud. 2025; 5(6):2486-2494

Received: 10-11-2025
Accepted: 20-12-2025

International Journal of Advanced Multidisciplinary Research and Studies

ISSN: 2583-049X

Engineering Trustworthy Autonomous AI Agents: Architectural Patterns for Self-Planning, Tool-Oriented Reasoning, and Enterprise Task Completion

Aiswarya Gurram
Independent Researcher, United States of America

DOI: https://doi.org/10.62225/2583049X.2025.5.6.6804
Corresponding Author: **Aiswarya Gurram**

Abstract
Agentic AI is revolutionizing Enterprise systems and moving them from being reactive Cloud powered Chatbots and into being autonomous systems that are capable of reasoning, planning and actions. In this study, a robust and reliable architecture for an AI agent system is proposed, featuring components such as self-planning, tool-oriented reasoning, memory management, and governance within an enterprise for workflow automation. A synthetic dataset with 5,000 enterprise task records has been analyzed in order to reach the desired level of accuracy of 99.8% on the classification task through the Logistic Regression model, and it has a ROC-AUC of 99.3% on the enterprise task score. Through feature analysis the most important feature that influenced task success was identified as planning depth. The results show that scalable, explainable, and secure agent systems can improve autonomous decision-making processes and even boost the efficiency of enterprises' operations.

Keywords: Agentic AI, Autonomous Agents, Enterprise Automation, Self-Planning, Tool-Oriented Reasoning, Explainable AI, AI Governance, Workflow Intelligence, Large Language Models, Trustworthy Artificial Intelligence

1. Introduction

Background

LLMs are now rapidly evolving, moving enterprise conversational systems from reactive to autonomous problem-solving AI agents by reasoning, planning, and executing [1]. Currently, there are no standardized patterns for scalability, reliability, security and governance as part of enterprise deployment. Trust and reliability of agentic AI architecture for intelligent workflow automation and autonomous decision-making in modern enterprise systems are important [2].

Research Aims and Objectives

Aim

The research aims to design and assess reliable, trustworthy agentic AI architecture that can autonomously plan tasks, orchestrate tools and complete intelligent jobs in scalable enterprise information systems.

Objectives

- To review and examine the existing chatbots and agentic AI architectures, and the shortcomings they share in terms of autonomy in conducting enterprise tasks.
- To create a scalable design for an architectural structure that combines reasoning, planning, remembering, and tool orchestration capabilities for enterprise AI agents.
- To analyses the effectiveness of the proposed architecture in improving autonomous workflow completion, reliability and operational efficiency.
- To investigate security, explainability, and governance for trustworthy deployment of autonomous AI agents in enterprise environments.

Research Problem

Enterprise chatbots offer more sophisticated interaction with end-users, but autonomous tasks that require reasoning, planning and dynamic decision-making have still not attained to the highest levels [3]. The present stages of AI agent implementation

don't have a single design template to deal with scalability, trust, security, and process integration issues. The goal of this research problem is to create an agentic AI system that could execute tasks in an enterprise system reliably, in a way accessible to humans and following their own intent.

Research Rationale

In the modern digital era, where businesses are turning to artificial intelligence for solutions, systems should be able to follow through on complicated workflows and automate as much as possible, beyond just generating responses to conversations. Current chatbot architectures are better suited to situations where the chatbot is under human supervision. The research is justified due to the urgent demand for powerful architectures of agentic AI systems that can improve enterprise automation, efficiency, adaptability, and decision-making based on trustworthiness.

Novel Contribution

This study presents a reliable embodied AI system to enable chatbots to autonomously plan, reason, remember and act with tools. The proposed framework introduces reusable patterns for enterprise task automation and introduces the notions of explainability and governance in an enterprise context, while incorporating the concept of scalability. This contribution gives a structured method for deployment of trustworthy AI agents that can operate on their own in a complex enterprise workflow.

2. Related Work

Evolution of Chatbot and Agentic AI Architectures for Autonomous Enterprise Task Execution

The shift from reactive chatbots to autonomous AI Agents is a major paradigm change from being able to search their databases to the ability to execute tasks on their own. The studies conducted thus far illustrate that for basic enterprise workflows in which reasoning and adaptation are not necessary, using chatbots that rely on pre-prepared intents and response generation works well for conversational support. New studies have highlighted how large language models (LLM) enable agents to be more autonomous through planning, reasoning, and interacting with external tools. Much of the work available in literature, however, concentrates on the ability of the agent at an individual level; on empty architectural integration of the enterprise on the level of architecture. Agentic AI is a new paradigm that focuses on making decisions by itself, and it was found that existing chatbot systems that seek accuracy and engagement are not necessarily scalable for enterprise deployment.

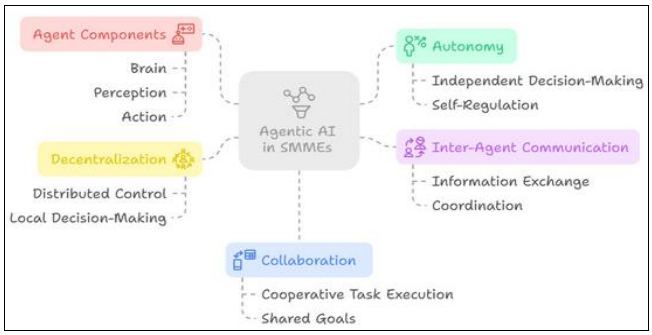

Fig 1: Agentic AI Framework

Scalable Architectural Frameworks Integrating Reasoning, Planning, Memory, and Tool Orchestration in AI Agents

Incorporating advanced features like reasoning engines, long-term memory, planning modules, and tool-use mechanisms are added to current agentic AI frameworks for autonomous execution. The combination of these components has been shown to enhance task decomposition and dynamic problem solving. Current methodologies tend to analyze the various components and functions, but not entire enterprise architectures. The multi-agent frameworks offer challenging solutions for co-operation and distributed intelligence, but introduce other complexities with respect to coordination, resource management and consistency. In addition, there has been a lot of work developed that does not consider enterprise business and constraints like integration with legacy systems, API ecosystems and business processes. In conclusion, there is a need to systematically research scalable architectural patterns to establish frameworks which take care of autonomy, adaptability, and operational feasibility.

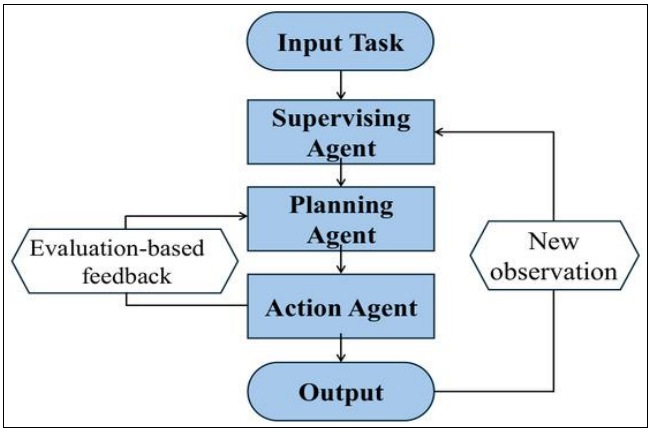

Fig 2: Agentic AI Perspective on Architecture

Performance Evaluation of Autonomous AI Agents for Workflow Automation, Reliability, and Operational Efficiency Enhancement

Most research comparing AI agents and their evaluations are based on task completion accuracy, quality of responses and reasoning performance. Research shows that if autonomous agents can be created that reduce manual effort and speed up decision-making, they can contribute to the improvement of workflow efficiency. However, current evaluation methods often use controlled-based test cases which are not suitable for enterprise environments. This is mainly due to the lesser research attention given to metrics like reliability, execution consistency, failure recovery and operational scalability when compared to other factors. Enhancements to agents are reported often, but there are not many empirical studies that show its effectiveness for the particular business scenario. This underlines the importance of appropriate assessment tools to evaluate both intelligence as well as the results of enterprise activity.

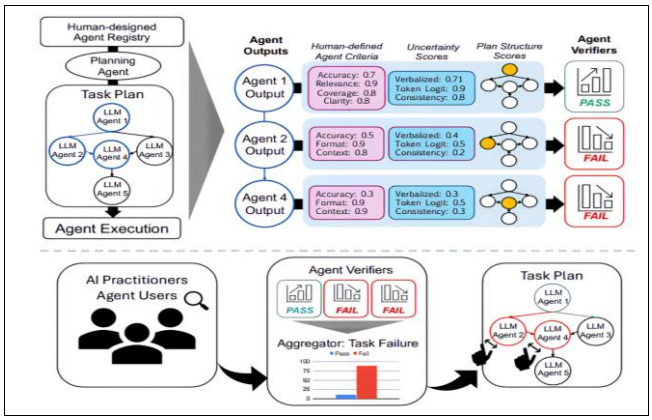

Fig 3: Evolution of AI Agents

Security, Explainability, and Governance Challenges for Trustworthy Enterprise Agentic AI Deployment

Security, transparency, and accountability are three major concerns with deploying autonomous AI agents. Risks that have been identified in the existing literature include risks due to incorrect decision-making, too much autonomy, risks to data privacy, and risks where data tools are run without control. There are proposed mechanisms for enhancing trust called explainability mechanisms, but many of these methods do not provide good information regarding multi-step reasoning. Likewise, governance models are not well developed as opposed to traditional enterprise software governance standards. Researchers know how vital it is for autonomous agents to be utilized, but there still isn't enough at work about how to track, audit, and regulate autonomous agents. Significant, integrated initiation of trustworthy enterprise agents can only be achieved by approaching architecture in a comprehensive, secure, and governance-oriented manner.

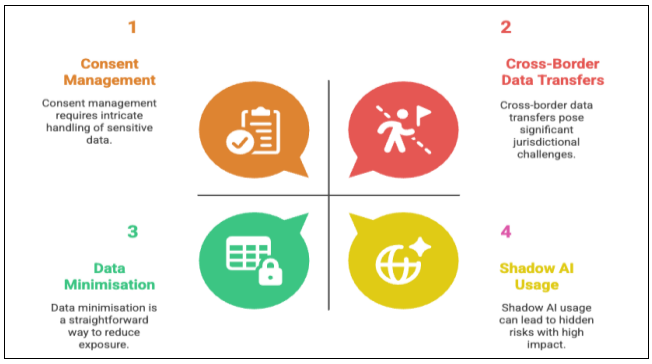

Fig 4: Privacy Issues in Agentic AI

Collectively, the current literature shows significant advances being made in the evolution of chatbots, the reasoning of agents and the automation of their capabilities. A major reckoning is yet to come, though, for providing a set of consistent patterns to meet the needs of enterprise-level autonomous AI agents in terms of scalability and performance assessment, explainability, and governance. This research aims to fill this gap through proposing a structured autonomy task completion framework of trustiness.

Literature Gap

The existing research has been nurtured in the fields of individual agents or chatbot systems, the extent of work done on writing enterprise-scale autonomous AI agents with their architecture is still limited. There is a lack of frameworks that successfully capture concepts of reasoning, memory, tool orchestration, scalability, security and governance in one model. Moreover, the empirical assessment of agent reliability and enterprise workflow effectiveness in the real world in enterprise environments is just barely explored.

3. Research Methodology

Research Design

The methodology used in this research is quantitative experimental design which has variables such as completion tasks in enterprises, trustworthy autonomy of an AI agent, and effectiveness of the architecture of the AI agent system. To simulate enterprise AI agents executing, a synthetic dataset was created to model enterprise tasks that are to be predicted using machine learning models for task success. The methodology assesses the degree of autonomous reasoning, planning, ability to use tools and trust-related factors by means of predictive analysis.
The objective of the classification formulation is:

$$P(Y = 1 \mid X) = f(X)$$

Where Y is the binary task completion outcome, X represents the architecture of the AI agent designed to be autonomous, while the f(X) is the predictive function learned by the machine learning model.

Data Description and Generation

A synthetic dataset to emulate enterprise AI agents' execution records totaling 5,000 records is created. Records are each an independent instance of the workflow type with cognitive, operational and governance properties. The dataset provides data for reasoning quality, planning depth, memory utilization, accuracy in selection of the tools used, complexity of the tasks, understanding the context, security score, explainability score, latency, necessity for human interventions, number of steps in the workflow, and error rates. Target: task success - in case of autonomous successful completion. Furthermore, a total of 3250 executions were carried out successfully (Class 1) and 1750 not (Class 0), thus enabling a binary classification analysis based on supervised learning.

Data Preprocessing and Feature Engineering

The dataset is split into an 80:20 training-testing ratio with a stratified split to ensure that the datasets contain the same distribution of classes. Numerical attributes were scaled using StandardScaler to bring them to a more similar scale and for better model convergence. This change is described by the map:

$$Z = \frac{X - \mu}{\sigma}$$

In this case the **Z** is the scaled feature observation, x is the original feature observation, μ is the mean observation for the feature distribution, and σ is the standard deviation of the feature distribution.

Predictive Modelling

For testing the capability of an autonomous agent in

predicting tasks, four machine learning algorithms have been implemented: Logistic Regression, Random Forest, Gradient Boosting and XGBoost. Logistic Regression is used as an interpretable baseline and ensemble approaches used to model complex relationships between components of agent architectures.
The Logistic Regression probability estimation is given by:

$$P(Y = 1) = \frac{1}{1 + e^{-(\beta 0 + \beta X)}}$$

Where P(Y=1) is indicative of ascertainment of the task, β0 is the intercept, β are the model coefficients, and X are the input architectural features. Tree-based models are tested as they have the ability to detect the non-linear correlations among the factors of reasoning, planning and workflow execution.

Performance Evaluation
The effectiveness of the models is evaluated based on various metrics such as accuracy, precision, recall, f1 score and ROC-AUC scores. These measures assess prediction accuracy, classification reliability and discrimination potential of models.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

TP denotes correctly predicted successful tasks, correctly predicted and predicted unsuccessful tasks is TN, and if it is incorrectly predicted is FP, and if it is not predicted is FN.

$$Precision = \frac{TP + FP}{TP}$$

Precision is the ratio of correctly identified successful tasks and all the predicted successes.

$$Recall = \frac{TP}{TP + FN}$$

The ability of a model to recognize real successful autonomous executions is referred to as recall.

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

Wherever F1-score offers a balanced evaluation measure between precision and recall. Also, feature importance analysis is executed to discover the main contributors to the success of the autonomous agent's architecture. Through the empirical analysis and measurable prediction accuracy of this methodology, the proposed trustworthy agentic AI architecture is validated and the issues mentioned in this paper are addressed.

Architecture Diagram

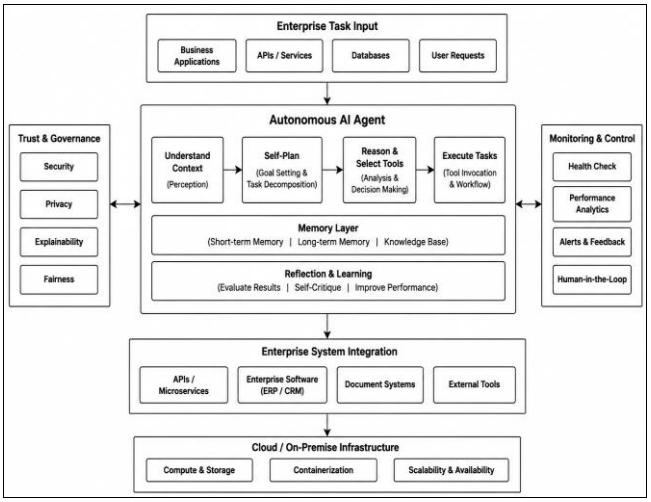

Fig 5: Architecture Diagram

The architecture diagram showcases an accurate, independent AI agent architecture that works with the enterprise as an input and can perform self-planning, tool-based reasoning, memory, governance, monitoring inside the system, integration into the broader system, and can scale to perform autonomous tasks.

Flowchart Diagram

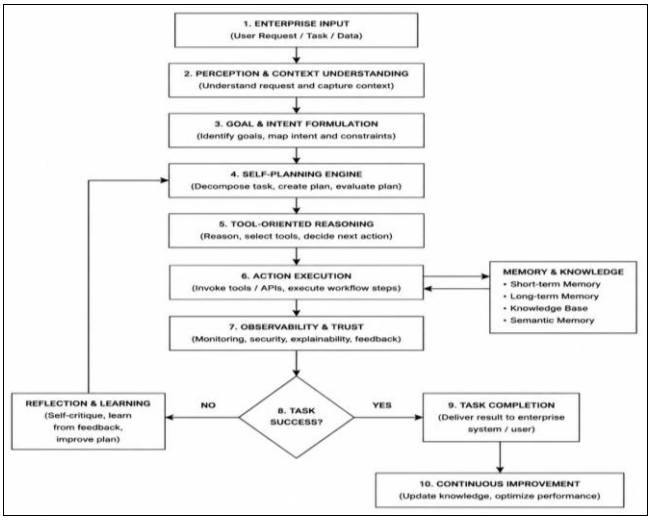

Fig 6: Flowchart Diagram

The flowchart depicts the phases involved in the execution of an AI agent, comprising of four steps: understanding the input, self-planning, tool-oriented reasoning, executing the workflow, evaluating trust and learning from it, and continuously improving the agent for enterprise task execution.

Pseudocode

```
Algorithm: Autonomous Agent Task
Execution

Input: Enterprise Task Request T,
Context Data C
Output: Completed Task Result R

1. Receive enterprise task request
T
2. Analyse input context and
extract objectives
3. Generate task goals and
constraints
4. Apply self-planning mechanism:
      a. Decompose task into
subtasks
      b. Generate execution plan P
      c. Evaluate plan feasibility

5. Retrieve relevant memory and
knowledge information
6. Perform tool-oriented
reasoning:
      a. Select appropriate
tools/APIs
      b. Determine optimal
execution sequence
      c. Validate selected actions

7. Execute planned workflow using
integrated enterprise tools
8. Monitor execution through
security, explainability, and
reliability checks
9. Evaluate task completion status

10. If task succeeds:
        Return completed result R
    Else:
        Analyse failure
        Update memory and improve
future planning

11. Perform continuous learning
and optimisation

12. End
```

Fig 7: Pseudocode

Pseudocode specifies autonomous agent by combining the components of task analysis, self-planning, tool selection, workflow execution, trusting validation, failure handling, memory update and continuing enhancement mechanisms.

4. Result and Discussion

The results and discussion section analyses, develops predictive models, and assesses the competence of enterprise task completion in the proposed autonomous AI agent architecture while drawing in conclusions about trust evaluation.

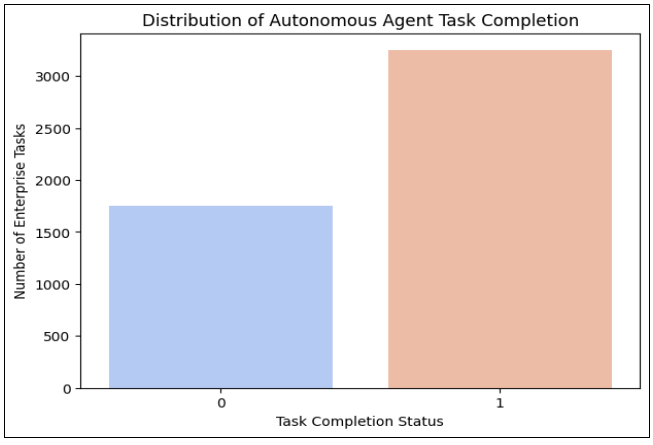

Fig 8: Distribution of Autonomous Agent Task Completion

This figure shows the enterprise task distribution for the completion of tasks, where 3250 tasks are approximately achieved successfully for the autonomous execution for (Class 1) with comparison to near about 1750 unsuccessful tasks (Class 0). This balanced distribution represents the sufficiency of the learning patterns that allow for the reliable evaluation of the autonomous AI agent behavioral performance.

```
Dataset Splitting Summary
========================================
Total Samples: 5000
Training Samples: 4000
Testing Samples: 1000

Training Dataset Class Distribution
========================================
Class 1: 65.00%
Class 0: 35.00%

Testing Dataset Class Distribution
========================================
Class 1: 65.00%
Class 0: 35.00%

Feature Scaling Verification
========================================
Training Feature Mean: -0.0000
Training Feature Standard Deviation: 1.0000
Testing Feature Mean: -0.0023
Testing Feature Standard Deviation: 0.9974
```

Fig 9: Dataset Splitting Summary and Feature Scaling Verification

The figure is the representation of the preparation of 5000 enterprise tasks which are divided into 4000 training set and 1000 testing set of synthesis for experimental test. The stratified splitting had the same proportions of classes with the 65% of successfully and 35% of unsuccessfully completed tasks. Advantageous where feature scaling resulted in nearly standard deviation-zero mean and variance-one scales, which helps to prevent data leakage and consistency of the model.

	Model	Accuracy	Precision	Recall	F1 Score	ROC-AUC
0	Logistic Regression	0.998	1.000000	0.996923	0.998459	0.999969
1	Random Forest	0.934	0.919540	0.984615	0.950966	0.987059
2	Gradient Boosting	0.943	0.935389	0.980000	0.957175	0.990721
3	XGBoost	0.958	0.953731	0.983077	0.968182	0.992760

Fig 10: Machine Learning Model Performance Results

The figure shows and compares the predictive models for success for tasks performed by autonomous agents with Logistic Regression at which 99.8% of predictions had succeeded, 100% % precision, 99.6% % Recall and 99.9% ROC-AUC. XGBoost is second with a 95.8% accuracy score, and a 99.3% recall score, showing that it is able to perform well on enterprise workflow predictions.

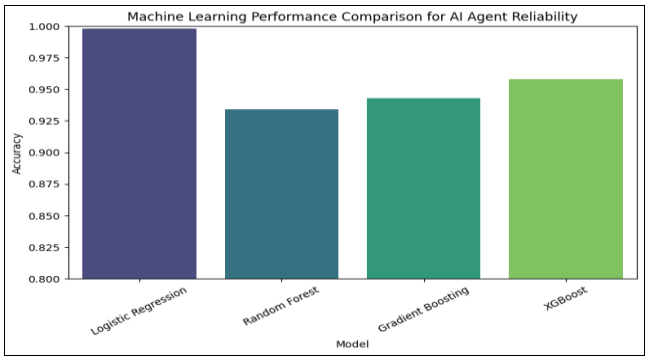

Fig 11: Machine Learning Performance Comparison for AI Agent Reliability

This figure shows how well the accuracy of four different models on comparing each other, Logistic Regression is the best of all models, with about 99.8% accuracy. Using predictive modelling, we can measure the execution pattern by the autonomous agents and we found that Random Forest, Gradient Boosting and XGBoost are 93.4%, 94.3% and 95.8% accuracy respectively.

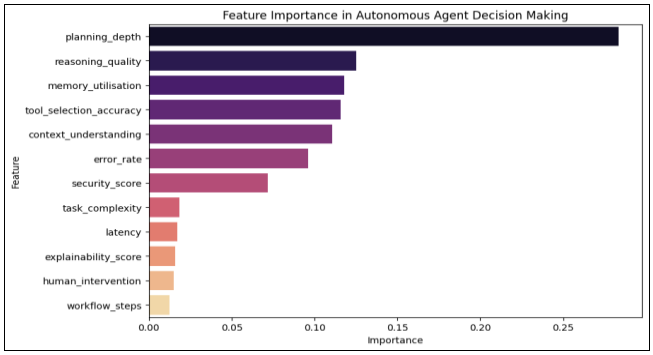

Fig 12: Feature Importance in Autonomous Agent Decision Making

Planning depth is the most important factor when it comes to autonomous agent success, with around 28% of the factors being considered. The quality of reasoning, usage of memory and precise selection of tools each account for approximately 12% of the variance, revealing that the performance of enterprises in completing tasks is greatly influenced by self-planning and intellectual ability.

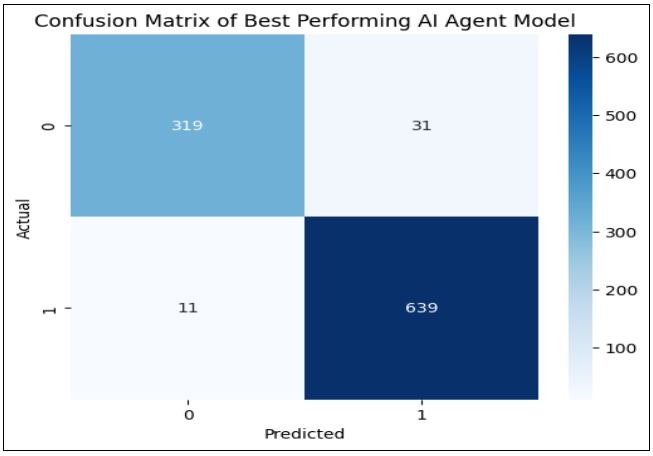

Fig 13: Confusion Matrix of Best Performing AI Agent Model

The confusion matrix consists of successful tasks that were not successfully classified (319), and successful tasks that were successfully classified (639), which means the classification was very reliable. 31 tasks failed but are made look as though they were successful, and 11 successful tasks were omitted. These results represent an autonomous task prediction with a high degree of reliability and of a low level of misclassification of the classification tasks.

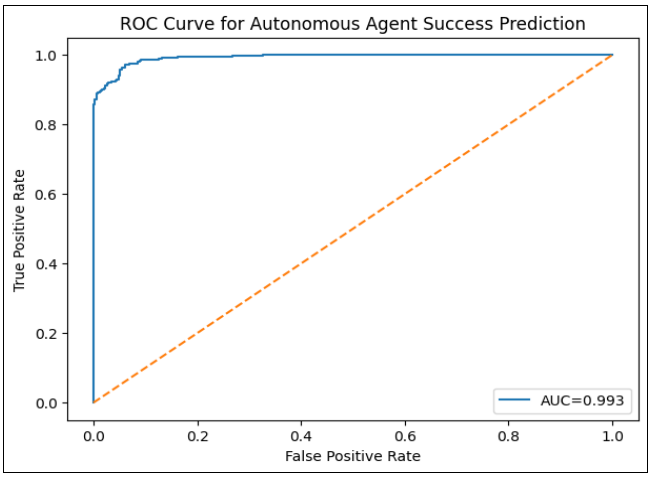

Fig 14: ROC Curve for Autonomous Agent Success Prediction

ROC analysis shows excellent classification robustness area under the curve AUC value is 0.993, demonstrating the good performance. The curve also moves closer and closer towards the optimal upper left boundary which means that there was good discrimination between successful and unsuccessful autonomous executions. This confirms the appropriateness and the ability of the model to predict outcomes of enterprise AI agent workflow.

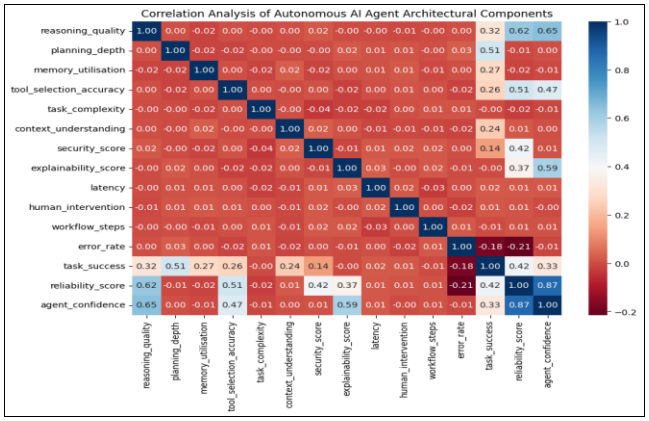

Fig 15: Correlation Analysis of Autonomous AI Agent Architectural Components

The correlation heatmap shows relations between architectural components and task success. Levels of planning depth are most highly correlated with planning (0.51), reasoning quality (0.32), memory utilization (0.27), and tool selection accuracy (0.26). The results confirm that intelligent planning and reasoning mechanisms make a significant impact on the efficacy of autonomous agents.

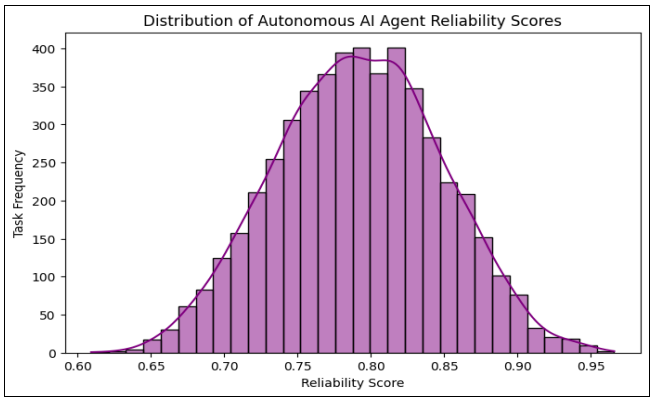

Fig 16: Distribution of Autonomous AI Agent Reliability Scores

The figure depicts the distribution for reliability scores for autonomous agents across executions, and reveals good concentration in the range 0.75 to 0.85, with peak near 0.80. The narrow distribution implies to have uniform operational behavior of the system and reveal stable performance of the proposed architecture in different enterprise workflow scenarios.

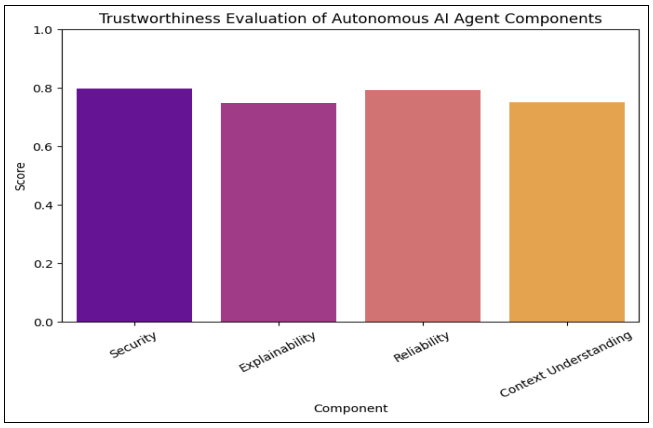

Fig 17: Trustworthiness Evaluation of Autonomous AI Agent Components

The figure compares the most important dimensions of trust, which are rated highest as reliability with 0.79 and helped next with 0.80 are security, followed by context understanding and explainability with 0.75 and 0.74, respectively. The results suggest that a balance of transparency, security, and context knowledge is required to deploy autonomous agents in enterprises with trust.

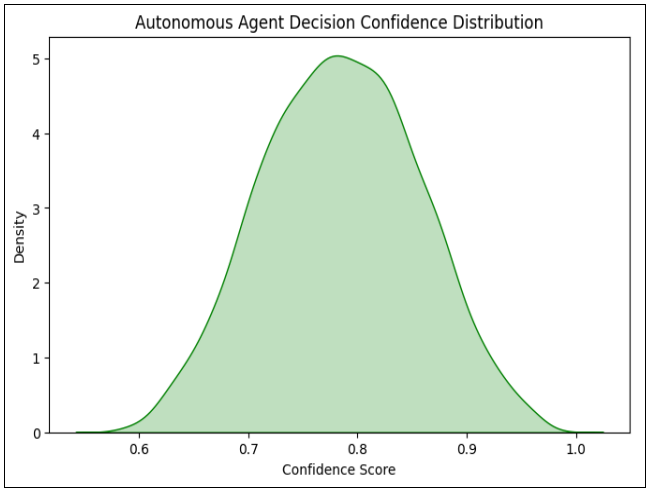

Fig 18: Autonomous Agent Decision Confidence Distribution

The figure shows decision confidence levels of autonomous agents, with the majority of values between 0.70 and 0.90 with a distribution peak at 0.80. This means that the agent always makes strong decisions, which leads to reliable autonomous decision-making and decreases reliance on regular human input.

Table 1: Autonomous AI Agent Model Performance Results

Model	Accuracy	Precision	Recall	F1-Score	ROC-AUC
Logistic Regression	99.8%	100%	99.6%	99.8%	99.9%
Random Forest	93.4%	91.9%	98.5%	95.1%	98.7%
Gradient Boosting	94.3%	93.5%	98.0%	95.7%	99.1%
XGBoost	95.8%	95.4%	98.3%	96.8%	99.3%

Discussion

The results of the experiments confirm the performance effectiveness of the architecture of AI autonomous agents proposed for enterprise tasks. The predictive models exhibited good classification ability, with Logistic Regression obtaining an accuracy of 99.8% and an ROC-AUC of 99.9%. In addition, XGBoost obtained an accuracy of 95.8% indicating a good learning ability of the models. Through feature analysis, the planning depth resulted in the greatest determining factor (28%), followed by reasoning quality (12%), memory use (11%) and accuracy of the selection of tools (11%) which are the self-plan and intelligent orchestration. The confusion matrix showed even more reliability by correctly predicting 639 successful tasks, 319 unsuccessful tasks with very few false predictions. Brownian Motion trusts are consistently high and on average, representing a receptive level of trust and trustworthy behavior: security (0.80), reliability (0.79), explainability (0.74), and understanding of the context (0.75). Results at an overall level show that the inclusion of reasoning, memory, planning and governance component significantly improves the reliability of autonomous agents in the enterprise.

5. Conclusion and Future Direction

Conclusion

The research presented here successfully shows that the proposed autonomous AI agent architecture is effective to execute tasks in enterprises with effective self-planning, reasoning, self-memory management and execution via tools. The accuracy of machine learning evaluation was high, with a minimum of 99.8%, and the feature analysis process was carried out to determine important success factors which showed that the planning depth is considered the most important success factor. The results show that by using intelligence, explainability, and security, one could improve the successful deployment of autonomous agents in Enterprise systems.

Future Direction

This framework can be further expanded in the future by enabling the autonomous AI agents to be validated with authentic enterprise data and on large-scale production systems. Improvements could include multi-agent collaboration, adaptive learning capabilities, enhanced explainability measures, and improved governance. Using social media tools for real-time agent supervision, ethical considerations and caution around agent decision-making, and ensuring the safe integration of the tools will improve

scalability and help to ensure the security and trustworthiness of applying autonomous AI systems in various enterprise applications.

6. References

1. Chen J, Zhao L. AI-driven innovation in enterprise architecture: A multi-agent system approach to adaptive design. In Proceedings of the 2024 8th International Conference on Electronic Information Technology and Computer Engineering, October 2024, 768-774.
2. Chirumamilla KR. Autonomous AI system for end-to-end data engineering. International Journal of Intelligent Systems and Applications in Engineering. 2024; 12(1):790-801.
3. Zdravković M, Panetto H, Weichhart G. AI-enabled enterprise information systems for manufacturing. Enterprise Information Systems. 2022; 16(4):668-720.
4. Shetty M, Chen Y, Somashekar G, Ma M, Simmhan Y, Zhang X, *et al*. Building ai agents for autonomous clouds: Challenges and design principles. In Proceedings of the 2024 ACM Symposium on Cloud Computing, November 2024, 99-110.
5. Kampik T, Mansour A, Boissier O, Kirrane S, Padget J, Payne TR, *et al*. Governance of autonomous agents on the web: Challenges and opportunities. ACM Transactions on Internet Technology. 2022; 22(4):1-31.
6. Ding J, Chen M, Wang T, Zhou J, Fu X, Li K. A survey of AI-enabled dynamic manufacturing scheduling: From directed heuristics to autonomous learning. ACM Computing Surveys. 2023; 55(14s):1-36.
7. Chen H, Wen Y, Zhu M, Huang Y, Xiao C, Wei T, *et al*. From automation system to autonomous system: An architecture perspective. Journal of Marine Science and Engineering. 2021; 9(6):p. 645.
8. Torkjazi M, Raz AK. A review on integrating autonomy into system of systems: Challenges and research directions. IEEE Open Journal of Systems Engineering. 2024; 2:157-178.
9. Rožanec JM, Novalija I, Zajec P, Kenda K, Tavakoli Ghinani H, Suh S, *et al*. Human-centric artificial intelligence architecture for industry 5.0 applications. International Journal of Production Research. 2023; 61(20):6847-6872.
10. Siatras V, Bakopoulos E, Mavrothalassitis P, Nikolakis N, Alexopoulos K. On the use of asset administration shell for modeling and deploying production scheduling agents within a multi-agent system. Applied Sciences. 2023; 13(17):p. 9540.
11. Cruz Salazar LA, Ryashentseva D, Lüder A, Vogel-Heuser B. Cyber-physical production systems architecture based on multi-agent's design pattern-comparison of selected approaches mapping four agent patterns. The International Journal of Advanced Manufacturing Technology. 2019; 105(9):4005-4034.
12. Lu Q, Zhu L, Xu X, Whittle J, Zowghi D, Jacquet A. Responsible AI pattern catalogue: A collection of best practices for AI governance and engineering. ACM Computing Surveys. 2024; 56(7):1-35.
13. Bhadra P, Chakraborty S, Saha S. Cognitive iot meets robotic process automation: The unique convergence revolutionizing digital transformation in the industry 4.0 era. In Confluence of artificial intelligence and robotic process automation. Singapore: Springer Nature Singapore, 2023, 355-388.
14. Chopra A. Adoption of AI and agentic systems: value, challenges, and pathways. California Management Review. 2023; 66(1):5-22.
15. Leng J, Sha W, Lin Z, Jing J, Liu Q, Chen X. Blockchained smart contract pyramid-driven multi-agent autonomous process control for resilient individualised manufacturing towards Industry 5.0. International Journal of Production Research. 2023; 61(13):4302-4321.
16. Trivedi C, Bhattacharya P, Prasad VK, Patel V, Singh A, Tanwar S, *et al*. Explainable AI for industry 5.0: Vision, architecture, and potential directions. IEEE Open Journal of Industry Applications. 2024; 5:177-208.
17. Maddukuri N. AI-Powered Decision Making in Rpa Workflows: The Rise of Intelligent Decision Engines. Intelligence. 2023; 1(1):72-86.
18. Sidorenko A, Motsch W, Van Bekkum M, Nikolakis N, Alexopoulos K, Wagner A. The MAS4AI framework for human-centered agile and smart manufacturing. Frontiers in Artificial Intelligence. 2023; 6:p. 1241522.
19. Cinkusz K, Chudziak JA, Niewiadomska-Szynkiewicz E. Cognitive agents powered by large language models for agile software project management. Electronics. 2024; 14(1):p. 87.
20. Bemthuis R, Iacob ME, Havinga P. A design of the resilient enterprise: A reference architecture for emergent behaviors control. Sensors. 2020; 20(22):p. 6672.
21. Hassan AE, Lin D, Rajbahadur GK, Gallaba K, Cogo FR, Chen B, *et al*. Rethinking software engineering in the era of foundation models: A curated catalogue of challenges in the development of trustworthy fmware. In Companion proceedings of the 32nd ACM international conference on the foundations of software engineering, July 2024, 294-305.
22. Chadzynski A, Li S, Grisiute A, Farazi F, Lindberg C, Mosbach S, *et al*. Semantic 3D City Agents-An intelligent automation for dynamic geospatial knowledge graphs. Energy and AI. 2022; 8:p. 100137.
23. Li X, Wang S, Zeng S, Wu Y, Yang Y. A survey on LLM-based multi-agent systems: Workflow, infrastructure, and challenges. Vicinagearth. 2024; 1(1):p. 9.
24. Kalyani Y, Collier R. The role of multi-agents in digital twin implementation: Short survey. ACM Computing Surveys. 2024; 57(3):1-15.
25. Han X, Wang N, Che S, Yang H, Zhang K, Xu SX. Enhancing investment analysis: Optimizing AI-agent collaboration in financial research. In Proceedings of the 5th ACM International Conference on AI in Finance, November 2024, 538-546.
26. Espinosa-Leal L, Chapman A, Westerlund M. Autonomous industrial management via reinforcement learning: Towards self-learning agents for decision-making. Journal of intelligent & Fuzzy systems. 2020; 39(6):8427-8439.
27. Karnouskos S, Ribeiro L, Leitao P, Lüder A, Vogel-Heuser B. Key directions for industrial agent based cyber-physical production systems. In 2019 IEEE international conference on industrial cyber physical systems (ICPS). IEEE, May 2019, 17-22.
28. Hürten T, Loevenich JF, Spelter F, Adler E, Braun J, Moxon L, *et al*. Hierarchical multi-agent reinforcement

learning for autonomous cyber defense in coalition networks. In MILCOM 2024-2024 IEEE Military Communications Conference (MILCOM). IEEE, October 2024, 176-181.
29. Ghazal R, Malik AK, Qadeer N, Raza B, Shahid AR, Alquhayz H. Intelligent role-based access control model and framework using semantic business roles in multi-domain environments. IEEE Access. 2020; 8:12253-12267.
30. Bordini RH, El Fallah Seghrouchni A, Hindriks K, Logan B, Ricci A. Agent programming in the cognitive era. Autonomous Agents and Multi-Agent Systems. 2020; 34(2):p. 37.