THE PENNSYLVANIA STATE UNIVERSITY
SCHREYER HONORS COLLEGE

Agentic AI in Business Processes

DHRUV MODI
Spring 2026

A thesis
submitted in partial fulfillment
of the requirements
for a baccalaureate degree of Enterprise Technology Integration
with honors in Information Sciences and Technology

Reviewed and approved* by the following:

David Fusco

Associate Teaching Professor and Director of Master's Programs

Thesis Supervisor

Carleen Maitland

Professor and Associate Dean for Research and Graduate Affairs

Thesis Honors Adviser

* Electronic approvals are on file.

Abstract

The rapid growth of Artificial Intelligence (AI) has led to the act of many people, groups, and organizations taking advantage of the unforeseen capabilities that this growing technology brings. More specifically, organizations and businesses have used AI to transition into a new form of business automation, which is led by the rise of automated systems that are capable of processing and perceiving data in order to make practical business decisions. This technology is called agentic AI, and it is one of the newest technologies that massive companies and organizations are looking to implement in order to make business processes more efficient. Unlike traditional machine-learning applications, agentic AI uses analytical decision-making paired with available data to curate decisions for different areas of operations within a business for both short-term and long-term decisions. This study investigates how secure and achievable the integration of agentic AI is for businesses in various areas of responsibility. It explores how agentic AI can prove to be useful when automating different processes and identifying various risks, while also showcasing how the risks of integrating agentic AI can be mitigated through following structured frameworks. This research will also cover how compliance, ethical, and technical issues can be major drawbacks for organizations looking to implement agentic AI into their processes, along with heavy cost considerations and abundant amounts of processed data that is required to be available. Ultimately, this study will contribute to the ongoing noise that artificial intelligence has brought to our daily lives, and will provide more information on how businesses can implement agentic AI in their systems in the most efficient way.

Acknowledgements **1**
Introduction **2**
Background & Context 2
Problem Statement 3
Research Objectives 4
Research Questions 5
Literature Review **6**
Evolution of AI in Business 6
Defining Agentic AI 7
Generative AI vs Agentic AI in an Enterprise Setting 8
Business Process Automation 9
AI Security and Governance 10
Methodology **11**
Research Approach 11
Data Sources 12
Analysis and Findings **14**
Use Cases of Agentic AI 14
Benefits of Agentic AI 16
Security Risks and Challenges 18
Governance Compliance Implementation 19
Discussion **21**
Key Findings 21
Implications for Enterprise Cybersecurity 21
Organizational Readiness 22
Cost Considerations 23
GenAI to Agentic AI Transition Costs 23
Agentic AI's Return on Investment 24
Agentic AI Failures 26
Conclusion **28**
Research Outcomes 28
Recommendations for Best Practice 29
Future Research 30
References **31**
Appendices **34**

Acknowledgements

I would like to express my thanks to Dr. David Fusco for being my thesis supervisor this past year. His guidance, support, and feedback were invaluable to this thesis and I truly appreciate him willing to take the time out of his busy schedule to work with me.

I would also like to thank Dr. Carleen Maitland for being my honors advisor. She is a huge reason why I was able to pick a meaningful thesis topic and answered the many questions I had throughout the process. Her guidance and feedback helped me succeed in completing this thesis tremendously.

Lastly, I would like to thank my family, friends, peers, and mentors for their support throughout my college experience. Their presence has helped me grow both academically and as a person, and I would not be the person I am today without them. I forever be grateful to have such a great supporting cast in life and will never take it for granted.

Introduction

Background & Context

Artificial Intelligence (AI) has recently made its way from a niche form of technology that could only be used in isolated scenarios, into a capability that modern organizations prioritize in order to gain a massive competitive advantage on their competition. The early AI deployments in businesses saw AI being used in a basic sense. Forecasting models, analytics, and automation of repetitive tasks were the main forte of AI uses for businesses at the time, and the results would ultimately be passed on for a human to make the decision using the data collected.

Newer advances in AI have rapidly taken over the world with their capabilities in adaptation, decision-making, and goal-driven responses for many large businesses to use for the benefit of their companies. The advancement of these new AI capabilities is known as agentic AI, and is able to do much more than create forecasts and automation. Interpreting context, planning actions, adapting to specific prompts, and collaborating with other components are new capabilities that agentic AI is able to do to help organizations who look to operate more efficiently while managing many different processes.

Looking at agentic from a business perspective, a massive shift from steady workflows to a faster-paced, reactive agent can be seen managing end-to-end workflows in different business areas. These agents can work in operations by monitoring inventory levels, planning logistics, and fulfilling procurement orders. They can also be useful in finance, where transactions can be monitored and contracts can be parsed through. Even in customer service, agentic AI can communicate with customers to solve issues and answer questions specific to customer needs or individual requests that are unique to each circumstance. These are just a few examples of how

agentic AI can play into driving business outcomes as an agent, rather than being a general support tool that isn't capable of automating tasks or making decisions.

Problem Statement

Agentic AI brings the idea of new capabilities for many organizations, but also unlocks new challenges for operations, security, and governance that are unprecedented from common approaches. While traditional workflows tend to revolve around controlled environments and anticipated results, agentic AI systems focus on autonomous and adaptive technology that integrate into core systems, but can largely increase risk for failures and cause issues for accountability (*Artificial Intelligence Risk Management Framework (AI RMF 1.0)*, 2023).

The security concerns revolving around agentic AI are especially prevalent, as agentic systems require access to an organization's sensitive data and tools in order to be used at its full capability. Different attacks, mainly prompt injections, can manipulate different applications into releasing sensitive data or performing harmful actions (Liu, 2023). In a broader sense, model performance isn't the only area of attention that AI should point to - it should also pay close attention to governance, measurements, and accessibility in order to be trustworthy and secure (*Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile*, 2024).

Additionally, organizations are facing immense pressure to ensure that proper governance and alignment for different AI deployments meet the required standards. According to the ISO/IEC 42001, the international management system for AI, representation for accountability, transparency, and risk management are expected to be implemented within the AI system's lifecycle (Javid & Welch, 2025). With this implementation, the process of conducting AI impact

assessments is also required for AI systems, as stakeholder accountability and legal compliance issues can be mitigated by this audit (Javid & Welch, 2025).

With these problems in mind, it is important to remember that the main question regarding this research is: How can organizations deploy agentic AI to automate business processes, while also delivering a positive return on investment and meeting both governance and compliance regulations. This thesis will focus on noting the capabilities of agentic AI, while identifying any drawbacks or risks that the tool may bring and evaluate different integration strategies that can meet governance expectations.

Research Objectives

This thesis will look to pursue these primary objectives:

1. Examine the use of agentic AI in business process automation

   This objective will identify who agentic AI agents can be used relative to business process automation and how it can be used towards automating complex, interdependent workflows.

2. Evaluate secure integration strategies for agentic AI

   This objective will dive deeper into how agentic AI agents can interact with enterprise systems securely, while paying close attention to key security concepts such as: access management, active monitoring, logging, and various security controls.

3. Assess implications for cybersecurity and governance

   Adding on the previous objective, this objective will dig into how agentic AI will impact a controlled environment, especially in the area of accountability, governance, and

compliance. It will consider how organizations can follow AI governance and risk-management structures when deploying AI agents in the system.

Research Questions

This thesis will be guided on answering these following questions:

1. How can agentic AI be used to automate complex workflows?

   This question will dive into how and where agentic AI can provide value to organizations beyond the traditional business process management systems.

2. What are the key security risks associated with agentic AI?

   This question will help address both new and existing risks that are relevant with the use of agentic AI, including: prompt injection, tool execution, data security, and access controls.

3. How can organizations implement AI while ensuring governance compliance?

   This question will evaluate how organizations can deploy agentic AI agents into the systems while meeting governance and compliance requirements for different systems.

Scope and Limitations

This thesis puts a focus on the implementation of agentic AI in business process automation, rather than the technicalities of the model itself. The main analysis of this research will be guided by observing three main objectives of agentic AI: business processes and automation, secure integration and risks, and governance and compliance methods to ensure proper deployment.

The scope for this research emphasizes the organizations that are medium-large in size, and points out the significant concerns of resilience, auditability, and data protection for these businesses. A key limitation to this research will be the lack of long-term evidence and common standardization for these technologies, as agentic AI is relatively new in the field and evolving quickly. Additionally, security threats and risks are also limited to the current standards in the field, as those features are constantly growing as agentic AI becomes more prevalent.

Literature Review

Evolution of AI in Business

Business automation has been around since before AI was what it is today, and revolved around procedural workflows defined by explicit rules and sequences. The early workflow standards deemed operating conditions and interpretable model interfaces as factors that resulted in automation success, and prioritized improvements in workflow engines, client applications, and monitoring applications to make this happen (Object Management Group, 2000). In the setting of a common enterprise, this type of approach is effective when there are structured inputs and low exception rates in workflows. This still doesn't account for the policy changes, unstructured content, and decisions that require human oversight, as these items can bring challenges when making a full-on automated workflow. Because of this, modern-day business process automation systems took the approaches that used data-driven support and monitoring to play a larger role in workflows ("Process Mining Manifesto," 2011).

Enterprise automation expanded by a huge margin when machine learning (ML) was introduced, as it trained systems to act based on detected patterns rather than following a set of

predetermined rules. For business actions, ML was able to use historical data to automate key decision points, such as - forecasting, routing, and prioritization - within the operations of the standard workflow. With the addition of deep learning, the automation for completing tasks with unstructured information, such as languages, audio, or images, improved drastically thanks to its capabilities of processing large amounts of data through neural networks (*Deep Learning Implementation in Enterprise*, 2022).

The introduction of language models (LMs) also proved to be a massive step in automating business processes by interacting with different tools in order to execute a task. More specifically, these LMs possessed the capabilities of calling different external tools, such as APIs, calculators, and other search features, to tackle the limitations of factual or arithmetic lookups (Schick et al., 2023). This use of LMs agentically through architectures that combine different actions such as planning, feedback loops, and tool summoning to gather information and complete objectives that require more than one step (Yao et al., 2022).

Defining Agentic AI

According to Gutowska and Stryker (2025), the definition of agentic AI refers to it mainly as an artificial intelligence that is able to accomplish a specific goal or task without the need of constant supervision (Gutowska & Stryker, 2025). Agentic AI is also composed of different ML models called AI agents that are able to solve problems in real time by mimicking decision-making capabilities of humans, and can also be used in multi-agent setups where individual agents each handle sub-tasks on their own to contribute to an overall goal (Gutowska & Stryker, 2025). Agentic AI also differs from traditional AI models in the sense that it focuses

more on adaptability rather than sticking with one process and working within predefined constraints.

Agentic AI capabilities can be applied to many real-world cases and can ultimately be used to automate many different business processes. Some examples of agentic AI in different scenarios, identified by Gutowska and Stryker (2025), include:

- Trading Bot: AI-powered bot is able to analyze stock prices and execute trades using patterns and predictive analysis (Gutowska & Stryker, 2025).
- Autonomous Vehicles: AI agents are able to pull data from GPS and sensors in order to improve safety and navigation (Gutowska & Stryker, 2025).
- Healthcare Agents: AI agents can pull data from patients and use a chatbot to provide treatment recommendations and feedback to clinics (Gutowska & Stryker, 2025).
- Cybersecurity Agents: AI agents can mitigate vulnerabilities by monitoring network traffic, logs, and behavior patterns (Gutowska & Stryker, 2025).
- Supply-Chain Agents: AI agents are able to place orders and adjust production levels in order to optimize inventory levels (Gutowska & Stryker, 2025).

Generative AI vs Agentic AI in an Enterprise Setting

In many corporate landscapes, generative AI (GenAI) tools have been present in an assistive way to help employees and executives to help amplify productivity at the base layer. A few examples of these actions can be writing a script, asking a question, or summarizing documents, which functions mostly as a support tool (Gutowska & Stryker, 2025).

Agentic AI brings more to an enterprise than an assistant, as its execution within structural systems can allow it execute decisions instead of informing them. These capabilities range from planning workflows, selecting tools, and adapting actions, which integrate into the optional core of enterprise systems. This integration does necessitate certain procedures for organizations in the governance and security areas however, as the risks from GenAI to Agentic AI can shift from output errors to deeper system errors (Amazon, 2025).

Business Process Automation

Business process automation (BPA) is traditionally implemented in a rule-based workflow along with robotic process automation, which both function in defined parameters where both sequences and inputs are specified in advance (Deloitte, 2025). The main issue with the traditional method is that each automation is built per use case, meaning additional workflows would require a new automation to be built (Deloitte, 2025). This limits the possibilities for scalability since these approaches use structured and static data while requiring human oversight (Deloitte, 2025).

The AI-driven BPA puts more of a reliance on multi-agent systems to orchestrate and automate different processes that can span over a number of departments. Amazon Web Services has taken this approach and created a project called Amazon Quick Automate, which handles multi-agent automation for different enterprise processes across departments, systems, UI and API interactions, and third-party systems (Amazon, 2025). By using natural language processing, function automations can be easily created from regular documentation, which overall ensures a balancing of autonomous AI with modern enterprise governance requirements (Amazon, 2025).

AI Security and Governance

Modern AI systems, including LLMs, tend to have tricky vulnerabilities since their layouts are non-deterministic and opaque. This makes it difficult to monitor, test, and analyze these systems, which can eventually lead to integrity failures and breaches (Scherlis, 2024). These weaknesses are also escalated in the sense that modern AI does not have a definitive separation between "code" and "data," which can pose a threat as processed data can function as instructions that create a huge vulnerability against injection attacks (Scherlis, 2024). Additionally, the areas for attackers to do damage is very broad, as they can enter through user prompts, inputted data, and even documents used to shape different models, which can be shaped in their favor. Risk mitigation for these types of vulnerabilities are also very system-reliant, as there is no consistent method for developers to identify or patch these doors of intrusion (Scherlis, 2024).

Data privacy and compliance measures are a huge part of AI security and governance, as AI systems depend on the collection, processing, and sharing of data, which can lead to conflict for consumer and organizational privacy. ISACA highlights how the "first line of defense" for AI governance is formed by privacy and cybersecurity, and that AI systems should comply with specific regulations such as that GDPR, CCPA, EU AI Act, and other AI laws in order to prevent unauthorized access and data manipulation (Demann, 2025). AI rules are also being reinforced by governments, as AI combined with privacy compliance obligations are being monitored with risks from third-party AI models, where vendors can use data to train models without consent, and the "right to be forgotten" regulation that requires AI systems trained on personal data to be deleted upon request (TrustArc, 2025).

Ethical AI practices are another huge point that organizations are looking to follow while developing agentic AI models. According to Short (2025), ethical AI refers to the approach of which AI is used in a philosophical way, focusing on the emotional impact of the use of AI (Short, 2025). Related to the term ethical AI is responsible AI, which puts more of an emphasis on how the AI is used in regards to compliance and transparency (Short, 2025). When building a framework, organizations use five principles: fairness, transparency, accountability, privacy, and security, to operationalize ethical AI practices (Short, 2025). Fairness accounts for ensuring that models use context as given and avoid tying in bias into outputs (Short, 2025). Transparency focuses on how the algorithm is built and what goes into it, while also watching out for bias through various entries – developers, data, or model designs (Short, 2025). Accountability assigns ownership and responsibility for the outcomes of the system, as organizations should be held accountable if a failure occurs in the system as the system itself cannot be held accountable (Short, 2025). Privacy verifies that personally identifiable information is protected and compliance with privacy laws are set within the system, with security doing the actions of encryption, access management, and anonymization to enable these principles within the system (Short, 2025).

Methodology

Research Approach

This thesis uses a qualitative approach when evaluating how agentic AI can be integrated into an enterprise's business process automation system, while adhering to organizational requirements for security, transparency, and governance. Due to the contemporary nature of agentic AI, long lasting data is very limited and has not been tested in every business scenario.

Because of this, this study revolves around the workflow analysis of agentic AI and its systematic mapping between agentic AI use cases and established frameworks.

Data Sources

This approach uses a collection of different online sources, each that applies to a portion, if not all, of the research conducted in this paper:

- Frameworks (Standard and Governance):
    - ISO/IEC 42001: AI Lifecycle Risk Management (Javid & Welch, 2025)
    - NIST AI Risk Management Framework 1.0 (*Artificial Intelligence Risk Management Framework (AI RMF 1.0)*, 2023)
    - NIST Cybersecurity Framework 2.0 (*The NIST Cybersecurity Framework (CSF) 2.0*, 2024)
    - NIST Generative AI Profile (*Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile*, 2024)
- Security and Risk
    - Prompt Injection attack against LLM-integrated Applications (Liu, 2023)
    - Weaknesses and Vulnerabilities in Modern AI (Scherlis, 2024)
    - The Actions of Language Models (Yao et al., 2022)
- Enterprise Workflow and Automation
    - Defining Agentic AI and its capabilities (Gutowska & Stryker, 2025)
    - Amazon's Quick Automate as an example for Agentic AI Automation (Amazon, 2025).
    - How AI Agents are Autonomising Business Processes (Deloitte, 2025)

    - Workflow/Process Management (Object Management Group, 2000) (“Process Mining Manifesto,” 2011)
    - Enterprise Deep Learning Implementation (*Deep Learning Implementation in Enterprise*, 2022)
- AI Privacy and Governance
    - AI and Data Privacy in the Age of AI: What’s Changing and How to Stay Ahead (TrustArc, 2025)
    - The New Triad of AI Governance: Privacy, Cybersecurity, and Legal (Demann, 2025)

Analysis and Findings

Use Cases of Agentic AI

Agentic AI provides the value of operating within different environments to execute goal-driven plans and manage complex tools while also adapting to various situations. This differs from traditional robotic process automation, which is less adaptable for common obstacles an enterprise faces on a daily basis. Agentic AI's versatility is highlighted by IBM, as they cover the different use cases for this technology in multiple fields:

- Front-End Services and Customer Experience
    - While regular chatbots use pre-coded scripts to handle customer interactions, agentic AI bots are able to gather responses from over time and use collected customer data to deliver real-time, contextual, and personalized responses during customer interactions. Beyond that, these agents are able to provide a cognitive conversation with the customer, and are able to gauge emotions within the conversation to assist the customer (Hayes & Downie, 2025).
- Financial Services
    - In the finance world, agentic AI is seen to "define a transformative era." The agent's capability to act quickly with large amounts of data shows great potential for financial services, as it can be used to improve decision-making, optimize workflows, and enhance compliance efforts. Part of this potential falls under audit actions, as enterprises can shift towards a "continuous audit" to monitor transactions or unusual patterns and resolve them to stay in line with compliance efforts (Hayes & Downie, 2025).
- Human Resources

    - Agentic AI agents are able significantly improve the human resources experience for the hiring process, onboarding, and HR requests from current employees. For the hiring process, these agents can complete many different tasks that consume the most amount of time, mainly being: analysing resumes, ranking candidates, and scheduling interviews. These capabilities can also be used for the onboarding process, as the AI agents are then able to create training plans and schedules for new employees to follow (Gutowska & Stryker, 2025). For regular HR requests, the AI agents can handle the actions of answering basic questions, as well as leave requests and training requirements (Gutowska & Stryker, 2025). IBM currently has an AI agent they use for HR requests called AskHR, which is able to automate over 80 HR requests, handle talent management, and standardize the hiring and promotion process (Gutowska & Stryker, 2025).
- Information Technology
    - AI agents in IT operations are able to perform many time consuming tasks for IT managers, including: managing infrastructure, optimizing system performance, and detecting unusual patterns. These agents are also able to assist developer teams and cybersecurity practices, as their continuous monitoring and threat detection allows them to automatically make necessary changes to systems or code when needed (Gutowska & Stryker, 2025).
- Supply Chain and Logistics
    - With the nature of supply chains being heavily decision-based, agentic AI systems are able to be extremely useful when adapting and making changes to inventory levels, procurement needs, and even transportation areas of fleeting and delivery

routes. These agents can also be used for forecasting, as they are able to analyze historical data and trends to provide both short-term and long-term advice for forecasting. In contract negotiations, the agents are capable of assessing different suppliers' metrics and numbers, which then allows them to provide insight for businesses to select the best contract for supply chain operations (Gutowska & Stryker, 2025).

Benefits of Agentic AI

- Process Optimization
    - Traditional AI systems tend to lack the ability of continuing tasks across different boundaries or environments. This is where agentic AI comes in, as these systems are designed to coordinate across different tools and environments due to its multi-agent layout, while also staying resilient to adaptations needed in the process in order to stay optimized. This move empowers AI agents to understand the bigger picture of organizational context and use dynamic actions to solve real-time changes (Deloitte, 2025). Additional services, such as Amazon Quick Automate, help add on to these environments and provide a centralized control service to reduce operational friction between different environments (Amazon, 2025).
- Cognitive Decision-Making
    - Agentic AI systems help carry the transition for traditional AI systems to go from basic generation to proactive outputs that follow a goal-oriented result (TechDemocracy, 2025). This is a much more efficient system that uses planning,

action, and evaluation to navigate business decisions, while traditional systems are only able to produce static outputs (TechDemocracy, 2025). Additionally, this autonomous system can allow AI to perform actions such like selecting tools, forming queries, and adapting to new information without the need for constant human intervention, which still enables the agent to complete an objective when dealt with shifts or incomplete information (Gutowska & Stryker, 2025).

- Time Savings
    - A big concern for many organizations when considering the implementation of agentic AI is the monetary and time investment that needs to be made initially. While this is true for the short-run time frame, the long-term benefits can prove to be useful for saving abundant time on future maintenance needs. According to Amazon (2025), traditional systems are known to be fragile, and can break when a minor change is made to the system. Issues like these can occur quite often, and require costs in both human oversight and allocated time to fix any issues. On the other hand, agentic AI systems are designed with “self-healing” capabilities from additional agents – saving organizations tons of maintenance and downtime costs that they would otherwise face from traditional systems (Amazon, 2025). Additionally, integrating a natural-language interface greatly reduces the cycle time during user interaction, as the interaction with a dialogue opposed to a manual navigating works quicker (Gutowska & Stryker, 2025). Overall, the integration of an agentic AI system prevents human stagnation actions, such as bottlenecks or handoffs from occurring, and allows for quicker actions within an organization.

Security Risks and Challenges

While agentic AI provides many new possibilities for organizations to increase their operational efficiency, it also institutes a whole new landscape for security risks and threats.

- Data Leakage
    - Since one of the main functions of agentic AI systems is to digest and route information across different enterprise systems, the possibility of data leakage is a massive risk. According to Scherlis (2024), this can occur from a "boundary collapse," where processed data can act as instructions and lead to confidentiality breaches, as mentioned earlier in this research (Scherlis, 2024). This type of vulnerability can really pose a problem in enterprise systems, as workflows that handle proprietary data can be negatively affected by a small leak or misfire (Scherlis, 2024). Additionally, the NIST AI RMF 1.0 notes "data leakage after decommissioning" as a constant lifecycle risk, as the termination of an agent's task extends the vulnerability to its memory and logs (*Artificial Intelligence Risk Management Framework (AI RMF 1.0)*, 2023). According to Amazon (2025), shifting towards continuous auditing and identity-based access controls can help mitigate any risk of data being exposed outside of the enterprise boundaries (Amazon, 2025).
- Attacks on AI Systems
    - In transitioning from generative AI to agentic AI tools, attackers tend to target the action layer of the system, focusing on the connectors and APIs that allow actual changes to be made (Amazon, 2025). As mentioned before in paper, agentic AI systems are prone to prompt injection attacks where malicious instructions are

inputted into a system to retrieve information. These attacks can lead to prompt theft, and execute unauthorized scripts to bypass other interface-level security and gain more private information (*Artificial Intelligence Risk Management Framework (AI RMF 1.0)*, 2023). The main solution to mitigating this risk is to build security around behavioral monitoring and low privilege, as there is no concrete way to "patch" this risk (Amazon, 2025).

- Compliance and Accountability Challenges
    - As mentioned earlier in this paper, the delegation of tasks from employees to AI systems can create a dead-zone in the accountability area, as systems and algorithms cannot be held responsible for its actions. A solution to this challenge is presented by the NIST AI RMF 1.0, as a functional framework of with the steps of governing, mapping, measuring, and managing can be used to document and validate controls throughout the agent's lifecycle (*Artificial Intelligence Risk Management Framework (AI RMF 1.0)*, 2023). Additionally, systems can enforce "step-points" within AI operations where a pause for human approval between phases is required before the next step takes place (TrustArc, 2025). The implementations can ensure that independent AI systems continue to follow the mandates, both legal and ethical, for the enterprise.

Governance Compliance Implementation

The implementation of agentic AI systems within an enterprise environment requires further care when focusing on the risk-management and compliance side of the integration. The agentic AI capabilities of autonomously conducting actions, making decisions, and using tools

command organizations to ensure that accountability, restrictions, and oversight actions are in place through the system's lifecycle (NIST AI RMF 1.0, 2023; Javid & Welch, 2025). This governance approach will require the implementation of the NIST AI Risk Management Framework (AI RMF 1.0) and ISO/IEC 42001. The AI RMF 1.0 framework will provide a functional layout of governance, mapping, and managing for a structure that identifies risks and validates controls through the lifecycle (*Artificial Intelligence Risk Management Framework (AI RMF 1.0)*, 2023). With this, the ISO/IEC 42001 will provide controls and ongoing commitment plans for continuously improving the autonomous operations for the broader system characteristics (Javid & Welch, 2025).

As mentioned previously in this paper, human accountability for agentic deployments is a key component for maintaining proper governance standards, as designated humans should be held responsible for their oversight on these systems, since the systems cannot be held responsible for malicious events (Short, 2025). Additionally, limiting the agent's access to specific APIs, environments, and data sets can help reduce the attack surface for lateral movements or credential abuse cases, which means that implementing least-privilege execution is a mandatory security and governance control for agentic systems (CyberArk, 2025; NIST CSF 2.0, 2024).

These findings suggest that continuous monitoring and visible auditability are essential governance requirements for agentic systems throughout its implementation and lifecycle. The logging of all actions and evaluation of system behavior play a key role in regards to this, as agentic system behaviors may adapt over time, meaning that persistent governance is required rather than static approval gates (*Artificial Intelligence Risk Management Framework (AI RMF 1.0)*, 2023). Adding on to this, impact assessments should also be conducted to evaluate

components like privacy, security, and transparency, to ensure that the system stays in line with ethical, technical, and legal requirements when it comes to governance standards (Javid & Welch, 2025; Demann, 2025. Overall, governance compliance is not only focused on preventing failures, but gauges the visibility of how the automation works and whether it stays up to par with reviewing and accountability requirements. By meeting these attributes, agentic AI systems can be used properly by organizations without being a legal, external, or internal liability (NIST AI RMF 1.0, 2023; NIST Generative AI Profile, 2024).

Discussion

Key Findings

Agentic AI's main value of interpreting context and coordinating with enterprise rules allows it to help reduce friction within organizational workflows, while traditional generative AI tools are only capable of producing static outputs. These strengths can be used in many different use cases as well, however, organizations must be willing to provide the proper technical and compliance efforts – specifically adopted from NIST AI RMF 1.0 and ISO/IEC 42001 – so that safety and transparency is prioritized over automation within the enterprise.

Implications for Enterprise Cybersecurity

The scope of cybersecurity for agentic AI shifts from focusing on access to focusing on actions. Within agentic models, systems are capable of making actual changes to the enterprise, compared to traditional AI models where outputs are required to pass through human approval.

Because of this, cybersecurity efforts must focus on not only preventing unauthorized access, but also unauthorized actions as well. As mentioned earlier in this research, the threat landscape would be expanded through a boundary collapse, in which attackers can look to penetrate the agent's toolchain in order to make systematic changes to their advantage (Scherlis, 2024). Because of this, security operations towards continuously monitoring behavior and applying "least privilege" for AI systems is essential for mitigating these vulnerabilities.

Organizational Readiness

A requirement for agentic AI systems that is underestimated, yet extremely significant, is organizational readiness and mature systems. This can include features such as:

- Data Maturity: having high-quality, processed data available for agents to consume
- Organizational Alignment: having coordination between different enterprise units
- Governance: enforcing proper monitoring and frameworks for impact assessments
- Financial Justification: a quantified cost-benefit analysis with a short-term or long-term goal, depending on the organization's needs

Ensuring that these features are present in medium to large enterprises helps reduce the risk of deployments for these organizations and verifies that enterprises are able to manage agentic AI agents.

Cost Considerations

The cost of implementing agentic AI systems can vary depending on the size of an organization and the level of sophistication desired, while also spanning across multiple stages. Additionally, this implementation requires a substantial investment upfront, and will have to cover the technical costs, such as hardware, software, storage, and governance efforts in monitoring, access management, and security infrastructure.

While these investments are not cheap, they are necessary in covering the possible vulnerabilities and risks that come with agentic AI, as security breaches and outdated technical systems can be a huge hurdle for improving efficiency and increasing productivity.

The main return on investment comes in the long-term scenario, as lower maintenance costs and downtime can save organizations from future costs and time commitments. However, these savings are not entirely guaranteed and depend heavily on how ready the organization is to implement agentic AI systems into their operations, as mentioned earlier in this research.

GenAI to Agentic AI Transition Costs

When discussed in the sense of transitioning from generative AI to agentic AI systems, organizations must account for structural labor and operational costs in order to automate decision-making and execution processes that agentic AI systems bring compared to GenAI systems (Deloitte, 2025). These drivers include:

- **Labor Cost Reduction**: Agentic AI's capability of automating certain processes can reduce up to 20% of a medium enterprise's manual coordination tasks, which can ultimately yield a substantial amount of recurring savings (Gutowska & Stryker, 2025).

- **Maintenance Resilience**: Agentic AI systems utilize “self-healing” capabilities that can command a more proactive approach that requires less human intervention. This approach helps companies shift IT costs from repairing maintenance requests to strengthening proactive approaches for these systems to allow them to self-heal, which also can reduce down-time and potential revenue loss (Amazon, 2025).

Agentic AI’s Return on Investment

Calculating the return on investment (ROI) for agentic AI systems can be a challenge for organizations, as there are many moving parts that need to be considered in order to provide an accurate cost estimate. According to Amazon Web Services (2026) and Monetizely (2025), there are a few key steps each organization should take in order to accurately estimate the overall return on investment agentic AI can provide for their systems.

The first step in evaluating the ROI is to define the current cost-structure of the organization. This benchmark will help determine the “human cost curve” opposed to the agentic performance, which requires organizations to quantify their labor costs, process inefficiencies, and compliance overheads (Amazon Web Services, 2026).

The second step for calculating the ROI is to account for an enterprise’s total cost of ownership. This includes the upfront cost – which is very high due to the depth of integration, as well as the ongoing monitoring, maintenance, and adaption capabilities of autonomous agentic systems. While the upfront costs are heavy for an organization to account for, the marginal cost for each additional transaction in the agentic system declines marginally, making it easier and more appealing to scale for larger organizations (Monetizely, 2025).

Following the calculation of the total cost of ownership, organizations should then aim to strategically plan out their cost reduction and revenue value creation methods in order to gauge the full outcome of this investment. For cost reduction metrics, organizations can evaluate components such as: labor savings, reduced error-related losses, and reduced compliance penalties to see how much their costs can be lowered by, as well as evaluate methods like quicker decision cycles, improved marginal accuracy, and higher transactional outputs in order to create value and revenue.

The autonomy levels of the agentic systems will come into play for this next step, as different agentic systems operate with different levels of freedom. This means that organizations will have to evaluate the ROI based on the agent's autonomy tier, which can be assistive, semi-autonomous, or fully autonomous (Amazon Web Services, 2026). Organizations should also include an expected risk cost in their modeling, so that they can acknowledge possible flawed executions by the system which have direct financial impact (Scherlis, 2024). With these calculations in place, organizations should calculate ROI using this formula:

$$ROI = \frac{(\text{Cost Savings} + \text{Revenue Gains} - \text{Expected Risk Cost} - \text{Total Cost of Ownership})}{\text{Total Cost of Ownership}}$$

In addition to the ROI formula, the expected risk cost can be calculated as follows:

$$\text{Expected Risk Cost} = P(\text{Incident}) \times \text{Financial Impact}$$

Since many agentic AI systems take time to process large amounts of data and continuously improve, the ROI should be evaluated over the span of 3-5 years. AWS stresses this, as the break-even point occurs when the cumulative savings transcend the initial implementation and governance costs, which happens when the transaction volume for the systems reaches a point where human scaling becomes substantially more expensive than having an agentic system in place (Amazon Web Services, 2026).

Agentic AI Failures

Despite frameworks discussed earlier in this research, real-world failure patterns suggest that these measures can prove to be insufficient in practice. While agentic AI systems provide many benefits beyond imagination for organizations, its autonomous actions and structural evolution can also create a larger surface for possible failures. These failures do not occur in a single event, as they typically occur in a few specific panic moments identified by CyberARK:

- The Crash (Operational Breakdown): This failure is normally caused by external dependencies, such as schema changes or model updates which can shift the AI's reasoning behavior and cause a loss of its "muscle memory," resulting in lower redundancy during system outages (Moss, 2025)
- The Hack (Credential/Identity Abuse): In this scenario, hackers can access the credentials of an AI agent, which is assumed to be a trusted worker in an organization, and use their privileges to move laterally or steal data. This can be huge problem, as attackers do not have to exploit a vulnerability and can mimic agent behavior, which could possibly give them more time in the system before they are detected (Moss, 2025).

- The Deviance (Failed Goal Alignment): This failure occurs when an agent functions properly, but prioritizes the wrong objective. This can happen when an agent faces a new environment, incorrect prompts, or poisoned data, and can really push an organization off of its course of action (Moss, 2025).

Another big technical driver for failure in agentic workflows is the propagation of errors in step of given tasks for an organization. These workflows, without the presence of human reviews, can cause two main dynamics which lead to system failures:

- Compounding Error Costs: This failure happens when a small error in a beginning step carries over to follow steps in a workflow. This ultimately causes the final output to be filled with errors as a result of the initial error, and requires more human labor to make up for the time the automation process was supposed to save (Chandarasekhar et al., 2025).
- Rule Brittleness: Many enterprise rules can be extremely exhaustive and could technically be bypassed in certain situations where the rules can be bent. This can be done efficiently by knowledgeable human employees who are able to bend the rules in their favor, but cannot be done by agent systems that follow the rules precisely. This can fail many companies in real-world scenarios, as contextual adaptation is necessary when knowing the rules properly (Chandarasekhar et al., 2025).

With these drivers in mind, both organizational and technical, it is important for organizations to be aware of the current agentic system failures occurring in different industries right now. Gartner (2025) has also conducted a forecast for these implementations and predicts

that 40% of agentic AI projects will be canceled by the end of 2027 due to these drivers and lack of value for organizations (Gartner, 2025). With this, roughly 93% of enterprises still intend to adopt AI agents into their system, even though in an experiment conducted by Debevoise & Plimpton (2025), the highest performing agent was only able to achieve a 2.5% success rate for full automated agents in organizations (Chandarasekhar et al., 2025). Overall, these failures are extremely important for considering agentic AI systems, as they stress the importance of organizational readiness and defined strategy to make the implementation successful.

Conclusion

Research Outcomes

Overall, this research has explored the integration of agentic AI within the modern enterprise, specifically focusing on autonomous process automation combined with governance and strategic decision-making. The main findings resulted in the conclusion that agentic AI provides a different, better, and more adaptable set of actions like interpreting context, utilizing tools, and generating plans without much human intervention.

Across various fields and use-cases, as discussed in section IV, agentic AI proves to greatly reduce organizational bottlenecks with its ability to “self-heal” and avoid failures in automation, which increases the likelihood of organizations to spend resources on future maintenance and stability fixtures. However, the research also identifies this autonomy as a proportional risk area, as data leakage, prompt injection, and accountability issues can be exploited through these tools. Additionally, this research found that agentic AI is heavily reliant on governance efforts, such as commitment to the NIST AI RMF 1.0, NIST Generative AI

Profile, and ISO/IEC 42001 are essential for mitigating “dead zones” and ensuring that operations utilizing AI systems comply with legal boundaries. Overall, agentic AI systems can only provide a meaningful return on investment when its efficiency gains are also accompanied by proper compliance, accountability, trust, and other governance controls that keep systems running for long-term operations.

Recommendations for Best Practice

This research has also demonstrated that the implementation of agentic AI can be a challenge depending on the field and structure of the organization. Because of this, organizations should adopt the following practices derived from this research in order to make the most out of agentic AI:

- **Establishing Boundaries:** organizations should define the limits for every deployment within the agency, while also categorizing the workflows by risk and implementing “stop points” for human oversight and approvals.
- **Least Privilege Execution:** Agentic AI systems should be used and treated as service accounts with certain privileges, but also restricted to specific APIs, database, and tools for direct tasks in order to mitigate movement in the event of a breach.
- **Behavior Monitoring:** Since the behavior of agentic AI is unpredictable and not controlled, continuous logging and auditability are important to keep track of these agents in the event of an incident, as uniform security is insufficient.
- **Organizational Readiness:** A massive takeaway from this research is that organizational readiness is one of the most important, yet overlooked, aspects of

implementing agentic AI, as the necessary data, organizational structure, and alignment is required for agentic AI to provide its full benefits and ultimately give the organization a full return on its investment.

Future Research

With the scope of AI continuously evolving and growing, further investigation is warranted for this type of technology in many different areas. The first component that needs further investigation is the settled-in evidence of agentic AI's performance. Since agentic AI is so new to the workspace, there aren't many, if any, long-term examples of agentic AI running its course in an organization's workflow. As time passes, research and monitoring of these systems should be conducted to analyze if agentic AI systems can impact organizations with solid evidence. Additionally, future studies should also monitor the hidden costs associated with agentic AI systems, such as monitoring, compliance, and auditing, so that organizations can have a reference for how much the total ownership costs of these systems can be for both short-term and long-term investments. For combating vulnerabilities and possible breaches, further experimentation of agentic AI agents in sandbox or testing environments should also be used to prevent breaches from occurring. These environments could also be used to evaluate how different agents interact with each other in a controlled setting, which can increase efficiency for organizations without the risk of damaging systems that are already implemented.

References

Amazon. (2025). *Amazon Quick Automate*. Amazon Web Services. https://aws.amazon.com/quicksuite/automate/?trk=8868fc11-8f91-4682-bfed-f9b1f356f95d&sc_channel=ps&trk=8868fc11-8f91-4682-bfed-f9b1f356f95d&sc_channel=ps&ef_id=Cj0KCQiAx8PKBhD1ARIsAKsmGbdxgVfovnsNEKPhSzRSCf07TFDWqxgO9FxHkSfmsWxWH8BYft5U188aAsPHEALw_wcB:G:s

Amazon Web Services. (2026). Measuring the success and ROI of agentic AI systems. Amazon Web Services. https://docs.aws.amazon.com/prescriptive-guidance/latest/agentic-ai-economics/measuring-success.html

Artificial Intelligence Risk Management Framework (AI RMF 1.0). (2023, January 1). NIST Technical Series Publications. https://nvlpubs.nist.gov/nistpubs/ai/nist.ai.100-1.pdf

Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile. (2024, July 25). NIST Technical Series Publications. https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.600-1.pdf

Chandarasekhar, C., Gesser, A., Levy, K., Kelly, M., Liss, J., & Bernabei, D. (2025, Nov 4). *Why Agentic AI Often Fails and the Enduring Value of Human Judgment*. Debevoise & Plimpton. https://www.debevoisedatablog.com/2025/11/04/why-agentic-ai-often-fails-and-the-enduring-value-of-human-judgment/

Deep learning implementation in enterprise. (2022, Nov 10). Intelligent Automation Network. https://www.intelligentautomation.network/intelligent-automation-ia-rpa/articles/deep-learning-implementation-in-enterprise

Deloitte. (2025, July). The rise of collaborative automation: How autonomous AI agents are redefining business processes. https://www.deloitte.com/content/dam/assets-zone3/us/en/docs/services/consulting/2025/generative-ai-agents-in-collaborative-automation.pdf

Demann, Y. (2025, March 17). *2025 Volume 6 The New Triad of AI Governance Privacy Cybersecurity and Legal*. ISACA. https://www.isaca.org/resources/news-and-trends/newsletters/atisaca/2025/volume-6/the-new-triad-of-ai-governance-privacy-cybersecurity-and-legal

Gartner. (2025, June 25). *Over 40% of Agentic AI Projects Will Be Canceled by End 2027*. Gartner. https://www.gartner.com/en/newsroom/press-releases/2025-06-25-gartner-predicts-over-40-percent-of-agentic-ai-projects-will-be-canceled-by-end-of-2027

Gutowska, A., & Stryker, C. (2025). *What is Agentic AI?* IBM. https://www.ibm.com/think/topics/agentic-ai

Javid, A., & Welch, A. (2025, May 13). *AI lifecycle risk management: ISO/IEC 42001:2023 for AI governance | Amazon Web Services*. AWS. https://aws.amazon.com/blogs/security/ai-lifecycle-risk-management-iso-iec-420012023-for-ai-governance/

Liu, Y. (2023, June 8). *[2306.05499] Prompt Injection attack against LLM-integrated Applications*. arXiv. https://arxiv.org/abs/2306.05499

Monetizely. (2025, Aug 30). *How to Calculate ROI for Agentic AI: A Guide to Measuring Autonomous Agent Value*. Monetizely.

https://www.getmonetizely.com/articles/how-to-calculate-roi-for-agentic-ai-a-guide-to-measuring-autonomous-agent-value

Moss, Y. (2025, Oct 31). *Crash. Hack. Deviate: Three AI agent failures every enterprise must prepare to face*. CyberARK. https://www.cyberark.com/resources/all-blog-posts/crash-hack-deviate-three-ai-agent-failures-every-enterprise-must-prepare-to-face

The NIST Cybersecurity Framework (CSF) 2.0. (2024, February 26). NIST Technical Series Publications. https://nvlpubs.nist.gov/nistpubs/CSWP/NIST.CSWP.29.pdf

Object Management Group. (2000, April). *Workflow Management Facility Specification, V1.2*. https://www.omg.org/spec/WfMF/1.2/PDF

Process Mining Manifesto. (2011, Oct 7). *Process Mining Manifesto*. https://www.tf-pm.org/upload/1580737614108.pdf

Scherlis, B. (2024, July 29). *Weaknesses and Vulnerabilities in Modern AI: Why Security and Safety Are so Challenging*. SEI Blog. https://www.sei.cmu.edu/blog/weaknesses-and-vulnerabilities-in-modern-ai-why-security-and-safety-are-so-challenging/

Schick, T., Dwivedi-Yu, J., Dessi, R., Raileanu, R., Lomeli, M., Zettlemoyer, L., Cancedda, N., & Scialom, T. (2023, February 9). *[2302.04761] Toolformer: Language Models Can Teach Themselves to Use Tools*. arXiv. https://arxiv.org/abs/2302.04761

TechDemocracy. (2025, June 13). *How Agentic AI is Shaping the Future of Smart Decision-Making*. TechDemocracy. https://www.techdemocracy.com/resources/agentic-in-decision-making-186

TrustArc. (2025). *AI and Data Privacy in the Age of AI: What's Changing and How to Stay Ahead*. TrustArc. https://trustarc.com/resource/data-privacy-age-ai-whats-changing/

Yao, S., Zhao, J., Yu, D., Du, N., Shafran, I., Narasimhan, K., & Cao, Y. (2022, October 6). *[2210.03629] ReAct: Synergizing Reasoning and Acting in Language Models*. arXiv. https://arxiv.org/abs/2210.03629

Appendices

Use Case	Framework(s) Applied	Analytical Intersection	Relevant Components	Implications
Customer Experience	NIST GenAI Profile	Shifting from static, predetermined scripts to real-life, dynamic conversations.	Safeguards against prompt injection, gauging customer emotions, and pulling real-time contextual data	These agents will require validation for inputs while monitoring behavior in order to prevent data leakage during the conversation.
Financial Services	ISO/IEC 42001 NIST AI RMF 1.0	Transitioning to a continuous audit and monitoring systems, rather than periodically.	Lifecycle risk management and governance functions	These agents require a strong documentation and accountability skillset to ensure integrity when automated decisions are made.
Human Resources	NIST AI RMF	Shifting to an onboarding system and employee management system with less human bias and	Risk mapping, designated human checkpoints, and automated talent management systems	Bias monitoring and human checkpoints will be necessary to ensure that fair hiring and employee

		more privacy.		promotions/demotions are handled without bias.
Information Technology	NIST CSF 2.0	Moving to autonomous system changes that can align with cybersecurity controls	Access management, Continuous logging and threat detection	IT agents must be capable of conducting real-time monitoring while operating under least-privilege operations.
Supply Chain	ISO/IEC 42001 NIST AI RMF 1.0	Adapting to procurement and inventory management decisions autonomously	Forecasting inventory and demand with accuracy and patterns.	These agents require structure to prevent disruptions during procurement stages, as well as human checkpoint to prevent errors being made when final orders are placed.