VILNIUS UNIVERSITY
BUSINESS SCHOOL

DEEPTECH ENTREPRENEURSHIP PROGRAMME

Kipras Daugirdas

THE FINAL MASTER'S THESIS

SCALING AI AGENT FRAMEWORKS: ENTREPRENEURIAL CHALLENGES AND OPPORTUNITIES IN THE NEXT WAVE OF DEEPTECH

Student ____________
(signature)

Supervisor ____________
(signature)

Eglė Terminė, Doc., Dr.
Name, surname, academic title, scientific degree

Vilnius, 2026

SUMMARY

VILNIUS UNIVERSITY
BUSINESS SCHOOL
DEEPTECH ENTREPRENEURSHIP PROGRAMME
STUDENT KIPRAS DAUGIRDAS
SCALING AI AGENT FRAMEWORKS: ENTREPRENEURIAL CHALLENGES AND OPPORTUNITIES IN THE NEXT WAVE OF DEEPTECH

Supervisor - Egle Terminė, Doc., Dr.
Thesis prepared - 2026, Vilnius
Thesis volume - 73 pages
Number of tables - 4
Number of figures - 6
Number of references - 50

This Master’s thesis investigates large language model (LLM)-based agent frameworks, which, despite rapid technical advancement, struggle to scale beyond pilot deployments in enterprise environments. While agentic systems demonstrate growing functional maturity, their adoption remains constrained by organizational, governance, and integration challenges.

The study examines how agent frameworks are positioned within emerging platform ecosystems and identifies the factors that most strongly limit enterprise adoption and scaling. The research focuses on adoption constraints rather than technological performance, addressing the conditions under which agentic systems transition from proof-of-concept implementations to sustained organizational use.

The research adopts a qualitative, exploratory design combining a systematic literature review with expert interviews involving creators, integrators, and adopters. Reflexive thematic analysis is used to integrate technical and organizational perspectives.

The findings show that enterprise adoption is shaped less by agent capability than by data readiness, governance requirements, human oversight needs, and unclear value realization. Rather than progressing toward full autonomy, agentic systems are most adopted in configurations of supervised autonomy corresponding to Level 2 AI adoption. This dynamic is conceptualized as a maturity-adoption bottleneck, where organizational constraints limit the translation of technical capabilities into scalable deployment.

SANTRAUKA

VILNIUS UNIVERSITETO
VERSLO MOKYKLA
AUKŠTŲJŲ TECHNOLOGIJŲ VERSLO PROGRAMA
STUDENTAS KIPRAS DAUGIRDAS
DIRBTINIO INTELEKTO AGENTŲ SISTEMŲ PLĖTRA: VERSLUMO IŠŠŪKIAI IR GALIMYBĖS NAUJOJE GILIOSIOS TECHNOLOGIJOS BANGOJE

Darbo vadovas - Egle Terminė, Doc., Dr.
Darbas parengtas - 2026m., Vilnius
Darbo apimtis - 73 puslapiai
Lentelių skaičius darbe - 4
Paveiksliukų skaičius darbe - 6
Literatūros ir šaltinių skaičius - 50

Šiame magistro darbe analizuojama, kodėl didžiųjų kalbos modelių (angl. large language models, LLM) pagrindu sukurti agentai, nepaisant sparčios techninės pažangos, sunkiai peržengia bandomųjų ar pilotinių diegimų ribas organizacinėje aplinkoje. Nors agentinės dirbtinio intelekto (DI) sistemos pasižymi augančiu funkciniu brandumu, jų praktinis taikymas įmonėse išlieka ribojamas organizacinių, valdymo ir integracijos iššūkių.

Darbo tikslas - išanalizuoti, kaip DI agentai pozicionuojami besiformuojančiose platforminėse ekosistemose; nustatyti pagrindinius veiksnius, ribojančius jų diegimą ir mastelio didinimą organizacijose; išgryninti sąlygas, leidžiančias agentinėms sistemoms pereiti nuo koncepcijos prie nuolatinio naudojimo. Tyrime dėmesys telkiamas ne vien į techninį agentų pajėgumą, o į organizacinius taikymo apribojimus.

Metodologiškai darbas grindžiamas kokybiniu tyriamuoju darbu, apimančiu sisteminę mokslinės literatūros analizę ir ekspertinius interviu su agentinių sprendimų kūrėjais, integratoriais bei organizacijų atstovais. Duomenų analizei taikoma refleksyvioji teminė analizė, leidžianti integruoti technines ir organizacines perspektyvas.

Tyrimo rezultatai rodo, kad agentinių sistemų diegimą organizacijose labiau lemia ne jų techninės galimybės, o duomenų parengtis, valdymo ir priežiūros reikalavimai, žmogaus įsitraukimo poreikis bei neapibrėžta vertės realizacija. Vietoje nuoseklaus judėjimo link visiškos autonomijos, agentinės sistemos dažniausiai stabilizuojasi prižiūrimos autonomijos konfigūracijoje, atitinkančioje antrojo lygio dirbtinio intelekto taikymą. Šis reiškinys

konceptualizuojamas kaip brandos-įsisavinimo ribojimas, kai organizaciniai apribojimai riboja techninių galimybių pavertimą plačiai masteliojamu taikymu.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

The accelerating digitalization of economic activity and increasing pressure on organizations to improve efficiency, adaptability, and resilience have intensified interest in artificial intelligence (AI) as a transformative business technology. AI is generally considered to be the capability of machines to perform tasks that traditionally require human intelligence. Within this field, the concept of the agent represents a fundamental analytical unit. An AI agent is an entity capable of perceiving its environment, persisting over time, adapting to change, and pursuing goals through autonomous action. From a traditional perspective, a rational agent chooses to complete tasks that maximize an expected performance measure, given available information and embedded knowledge (Russell & Norvig, 2021).

Recent technological progress in Large Language Models (LLMs) has significantly expanded the practical relevance of agent-based systems. LLM-enabled agents are no longer limited to predictions but are increasingly capable of reasoning, planning, maintaining memory, and interacting with digital environments. These developments have shifted AI toward agentic systems capable of executing sequences of actions and participating directly in organizational workflows (Tang et al., 2024). As a result, AI agents are increasingly positioned not merely as support tools, but as semi-autonomous mechanisms used within business processes.

To conceptualize the usage of these agents within organizations, researchers and practitioners frequently apply maturity-based frameworks that distinguish between assistive, collaborative, and autonomous forms of AI integration. This thesis focuses on Level 2 AI Adoption, often referred to as Agent AI or Co-Worker AI, in which AI systems assume responsibility for specific business tasks while operating under human supervision (Hang & Chen, 2022). At this level, AI agents move beyond experimental use and begin executing defined organizational roles, marking a critical transition from pilot projects to operational deployment.

Despite the rapid technological progress, the usage of AI agent frameworks in organizations remains uneven. While many organizations demonstrate promising pilot implementations, few have succeeded in reliably utilizing AI agents in their core business operations. This discrepancy reveals a persistent gap between the technical capabilities of AI agents and adoption within organizations. In practice, challenges related to data readiness, system integration, system management, trust, and dependency on digital platforms frequently limit the realization of agentic AI beyond isolated use cases. These challenges are particularly evident for entrepreneurial and DeepTech ventures, which have limited resources and operate within platform-dominated ecosystems.

The research problem addressed in this thesis stems from the gap between the growing technical maturity of AI agent frameworks and organizations' limited ability to scale these systems to stable Level 2 adoption. While the existing literature extensively discusses the technical potential of agentic AI, there is limited empirical understanding of how entrepreneurial ventures navigate organizational readiness, governance requirements, and platform dependence when using AI agents in business operations.

The research objective is to examine the process of scaling AI agent frameworks within platform-based organizational environments.

The research aims to empirically examine and conceptually model how entrepreneurial ventures scale AI agent frameworks toward Level 2 AI adoption by managing technical, organizational, and platform-related constraints.

To achieve this goal, the thesis pursues the following interconnected **research tasks**: to systematize the theoretical foundations of AI agents, organizational adoption levels, dynamic capabilities, and platform ecosystem governance; to identify the principal organizational and technical barriers that hinder the scaling of AI agent frameworks beyond pilot implementations; to analyze the roles and interactions of key ecosystem actors involved in the creation, integration, and adoption of AI agents; to empirically examine how firms apply dynamic capabilities to bridge technical maturity and business adoption; and to refine a conceptual model explaining the transition toward Level 2 AI adoption in DeepTech entrepreneurial contexts.

The research adopts a qualitative, exploratory **methodology** underpinned by an interpretivist perspective. Empirical data are collected through semi-structured expert interviews conducted with representatives of different ecosystem roles involved in the ecosystem of AI agent development, integration, and adoption. The data are analyzed using qualitative thematic analysis, allowing patterns and mechanisms relevant to AI agent scaling to emerge from a practical point of view. The study further employs analytical triangulation by complementing empirical findings to established theoretical frameworks to enhance interpretive accuracy.

The structure of the thesis follows the methodological guidelines of Vilnius University Business School. Following this introduction, Chapter 2 develops the theoretical framework and presents the conceptual model, while Chapter 3 outlines and justifies the research methodology. Chapter 4 presents the empirical findings, which are analyzed and interpreted in Chapter 5. The thesis concludes with conclusions and practical recommendations. While the study is subject to limitations related to sample size and the rapid evolution of AI technologies, it aims to provide analytically backed insights into the organizational challenges of scaling agentic AI in today’s business environments.

1. THEORETICAL AND CONTEXTUAL FRAMEWORK

1.1. Artificial Intelligence and Agentic Systems

This chapter develops the theoretical and contextual framework that underpins the empirical analysis of AI agent scaling. Rather than presenting abstract theory in isolation, the chapter integrates foundational perspectives from artificial intelligence with contextual lenses drawn from DeepTech entrepreneurship, platform ecosystems, and strategic management. This structure reflects the theory that scaling AI agent frameworks toward Level 2 adoption emerges from the interaction between technical capability, organizational readiness, and ecosystem constraints.

The first theoretical pillar concerns the rapid advancement of Large Language Models (LLMs) and their role in enabling a new generation of AI agents. Recent studies demonstrate that LLM-based agents differ fundamentally from earlier predictive or rule-based systems, as they are capable of reasoning, planning, and interacting with users through natural language interfaces (Tang et al., 2025). These capabilities allow agents to operate as semi-autonomous entities rather than passive analytical tools. However, while such agents showcase impressive performance in controlled environments, deploying them in organizational settings introduces new integration challenges. In particular, their ability to perceive contextual information and execute multi-step workflows mandates that companies reconsider existing software architectures, data pipelines, and operational responsibilities (Tupe & Thube, 2025). This literature, therefore, emphasizes technical advancement but offers limited guidance on how organizations can integrate these changes to facilitate the stable operations of Level 2 AI systems.

The second theoretical pillar addresses this organizational challenge while using the Dynamic Capabilities framework. Prior research on DeepTech entrepreneurship suggests that technical superiority alone is insufficient for achieving sustainable scaling, particularly in high-velocity, uncertain environments. Organizations must also have the capability to sense emerging opportunities, seize them through appropriate business models, and continuously transform their resource base in response to technological progress (Teece, 2007; Warner & Wäger, 2019). This thesis adopts Dynamic Capabilities as its primary strategic management framework because it focuses on businesses' ability to adapt for adaptation in the face of uncertainty. However, much of the existing literature remains abstract, often treating digital technologies as generic resources and not as systems that introduce new forms of organizational risk, responsibility, and control. By applying this framework specifically to AI agent deployment, this research seeks to refine how dynamic capabilities operate when decision-making authority is handed over to semi-autonomous systems.

The third theoretical section recognizes that AI agents do not operate in isolation, but are used within broader ecosystems of digital platforms. AI agent frameworks are inherently dependent on foundation models, cloud infrastructure, and API-based integration environments, which are typically controlled by large platform providers (Luitse, 2024). The platform ecosystem theory, particularly as articulated by Hein et al. (2020), provides valuable material for analyzing how platform integrity and complementor generativity are balanced through governance mechanisms. The literature explains how ecosystems enable innovation at scale and reveals structural constraints that limit entrepreneurial freedom. When talking about AI agents, these constraints are especially visible, with platform governance decisions potentially directly limiting access to data, automation privileges, and deployment scope. However, existing studies tend to analyze platform dynamics at a generic level, without fully accounting for the heightened integration and control requirements associated with AI agentic systems.

Together, these three perspectives form a multi-layered analytical framework that links micro-level agent behavior, firm-level capability development, and ecosystem-level governance structures. A critical dimension that emerges across these sources of literature is the degree of AI implementation maturity. Prior research classifies AI agents along their capability maturity stages, ranging from basic command execution to fully adaptive and socially interactive systems (Tang et al., 2025). While such classifications are functional for technical comparison, they often overlook organizational trust, accountability, and role substitution, which are key factors that are used when businesses decide to utilize AI agents. Thus, this thesis uses a more constrained interpretation of AI maturity that considers organizational readiness alongside technical capability.

Attention is given to Level 2 AI adoption, which Sohn (2024) refers to as “AI Equalization” or “Agent AI.” At this stage, AI systems can execute specific business tasks independently, functioning as semi-autonomous collaborators rather than simple decision-support tools (Level 1) or fully autonomous organizational actors (Level 3). Although commercialization typically begins at Level 1, the empirical and theoretical evidence suggests that the most significant scaling challenges emerge at Level 2. This is the point at which firms must confront not only technical reliability issues, but also governance, liability, and control concerns that determine whether agentic systems can be embedded into daily operations.

From an entrepreneurial perspective, Level 2 adoption represents the first critical inflection point in the commercialization process. Enabling AI agents to perform tasks independently requires ventures to align technological development with organizational structures and ecosystem constraints. This theoretical chapter, therefore, combines insights from AI agent research, strategic management, and platform governance to establish an integrated foundation for analyzing how AI agent frameworks can be scaled by organizations.

By doing so, it provides the conceptual basis for the subsequent empirical investigation into how entrepreneurs navigate the relationship between technological advancement, organizational capabilities, and infrastructural oversight.

To ground this analysis, briefly review the foundations and evolution of artificial intelligence as a discipline, as these developments explain why contemporary AI systems differ fundamentally from earlier decision-support technologies. AI research has progressed through several distinct stages, moving from symbolic and rule-based systems toward data-driven models and, more recently, toward agent-based architectures capable of autonomous action. This technological progression is key to understanding why contemporary AI systems are completely different compared to earlier tools that only supported decisions and not made them.

The recent progress of LLMs has enabled AI agents to move from static predictions toward goal-oriented behavior, including planning, reasoning, and interaction with complex environments (Tang et al., 2025). Unlike traditional machine learning models, which generate outputs in response to predefined inputs, AI agents are designed to perceive contextual information and execute sequences of actions to achieve specific objectives. This shift marks a qualitative change in how AI systems can be deployed within organizational settings.

Agent-based systems demonstrate strong theoretical capabilities, their transition into real-world business environments remains constrained by integration, reliability, and control challenges (Zhan et al., 2024). As a result, the evolution of artificial intelligence must be understood not only as a technical trajectory, but also as a foundation for examining how increasingly autonomous systems can be scaled and governed in commercial contexts.

Before examining these governance and scaling implications in detail, it is necessary to clarify what is meant by artificial intelligence itself and how different conceptualizations of intelligence have shaped the development of agent-based systems. This conceptual grounding underpins the subsequent analysis of AI agents, autonomy levels, and their implications for DeepTech entrepreneurship.

Russell and Norvig (2021) systematize these debates by distinguishing four approaches to defining artificial intelligence: acting humanly, thinking humanly, thinking rationally, and acting rationally. The acting humanly approach exemplified by the Turing Test, which evaluates intelligence based on a machine’s ability to produce responses indistinguishable from those of a human. While historically influential, this approach focuses on imitation rather than functional effectiveness and provides limited guidance for designing systems intended to operate autonomously in organizational contexts.

The thinking humanly approach shifts attention toward cognitive modeling, emphasizing whether a system’s internal reasoning processes resemble those of the human mind. Although valuable for understanding human cognition, this perspective has proven difficult to operationalize, as human reasoning is neither fully observable nor consistently rational. As a result, its applicability to scalable AI systems remains limited, particularly in commercial environments where performance and reliability take precedence over cognitive similarity.

The rational thinking rationally approach, described as in “laws of thought,” seeks to formalize intelligence through logical reasoning. However, this approach encounters practical limitations when addressing real-world problems, as informal knowledge is difficult to formalize and computational complexity often renders perfect inference infeasible. These limitations reduce its relevance for AI systems expected to operate under uncertainty and time constraints.

Approach has become the dominant foundation of modern AI research. From this perspective, an intelligent system is defined as a rational agent that selects actions expected to maximize goal achievement given available information and environmental constraints (Russell & Norvig, 2021). Importantly, rational action does not require perfect logical reasoning; even simple reflexive behaviors can be considered sane if they contribute to favorable outcomes. This interpretation underpins the Standard Model of AI, in which agents are designed to optimize expected utility with respect to human-defined objectives.

From the perspective of this thesis, the rational agent framework is particularly significant because it establishes artificial intelligence as an operational construct rather than a purely analytical or representational one. By defining intelligence in terms of goal-oriented action under conditions of uncertainty, the acting rationally approach provides the conceptual foundation for examining AI systems as entities capable of executing tasks and influencing outcomes within real organizational settings. This interpretation is essential for understanding how AI agents transition from technical artifacts into economically relevant actors embedded in business processes and platform ecosystems.

Agent model did not emerge in isolation. It is the result of a broader historical evolution in how intelligence has been conceptualized, modeled, and implemented in computational systems. To situate contemporary AI agents within this trajectory - and to clarify why agentic systems represent a qualitative shift from earlier decision-support technologies - it is necessary to review the major paradigms that have shaped the development of artificial intelligence as a discipline.

AI research, particularly during its inception and formative years (1943-1956), was dominated by symbolic logic and rule-based reasoning. The Physical Symbol System Hypothesis proposed that intelligent behavior could be achieved through the manipulation of symbolic representations, enabling early systems to solve logical puzzles and prove mathematical theorems (Russell & Norvig, 2021). While these approaches demonstrated theoretical feasibility, they proved inadequate for real-world applications because rigid symbolic rules could not account for uncertainty, incomplete information, or environmental variability.

In response to these limitations, the field shifted during the 1960s and 1970s toward expert systems that relied on domain-specific knowledge encoded by human experts. Rather than general problem-solving algorithms, these systems focused on replicating expert decision-making within narrowly defined contexts. Although expert systems achieved technical success in controlled domains, their lack of adaptability limited their scalability. They struggled with ambiguous inputs, novel situations, and data outside predefined rule sets, ultimately constraining their commercial viability (Hang & Chen, 2022).

From the late 1980s onward, artificial intelligence adopted a more probabilistic and statistically grounded approach. Bayesian networks enable systems to reason under uncertainty by modeling probabilistic dependencies among variables, thereby supporting more robust decision-making in noisy environments (Weber et al., 2012). This shift laid the foundation for the resurgence of neural networks and the rapid expansion of deep learning in the 2010s. Advances in computational power, the availability of large-scale data, and the proliferation of Internet of Things (IoT) devices have enabled deep learning models to outperform human benchmarks in domains such as image recognition and strategic gameplay, establishing AI as a core driver of the digital economy (Hang & Chen, 2022).

Large Language Models (LLMs), while highly effective in processing and generating natural language, operate primarily as probabilistic sequence predictors without intrinsic agency or long-term autonomy (Tang et al., 2025). Their limitations in maintaining persistent goals or interacting with external environments motivated the development of agentic systems, in which LLMs function as cognitive cores embedded within broader architectures. Those architectures incorporate memory for contextual continuity, planning mechanisms for task decomposition, and tool-use capabilities for interacting with external systems and APIs (Tang et al., 2025).

Predictive models to agentic systems represent a critical evolutionary step toward operational autonomy. By enabling AI systems to execute sequences of actions and perform business functions independently, agent-based architectures support what Sohn (2024) describes as AI equalization. From the perspective of this thesis, this evolution is significant because it marks the point at which AI systems shift from analytical support tools to semi-autonomous actors within organizational processes. Understanding this historical trajectory provides the necessary foundation for distinguishing between traditional AI systems and contemporary agent frameworks, which is explored further in the following section through an examination of their architectural differences.

1.2. Intelligent Agents: Classical and LLM-Based Approaches

The concept of intelligent agents has evolved from strictly defined logical entities toward adaptive, probabilistic systems capable of operating in dynamic environments. In classical artificial intelligence, agents were primarily modeled as rational decision-makers that perceive their environment and select actions according to predefined rules or optimization criteria (Russell & Norvig, 2021). While these architectures provide a clear theoretical foundation, their applicability is limited in complex real-world settings characterized by uncertainty and incomplete information. It is also important to note that the term agent-based is used beyond LLM-driven systems; for instance, domain-specific decision and simulation frameworks employ agent-based approaches to model or optimize complex processes, such as prioritizing building retrofits (Lari et al., 2025). This distinction is relevant for this thesis because it separates agent-based modeling and optimization traditions from contemporary LLM-based operational agents embedded in business workflows.

Recent advances in generative artificial intelligence, particularly Large Language Models (LLMs), have enabled a new class of agent architectures with enhanced flexibility and contextual awareness. LLM-based agents differ from classical agents in that they rely on learned representations rather than explicitly programmed rules, allowing them to reason, plan, and interact across a wide range of tasks (Tang et al., 2025). This shift has significantly expanded the functional scope of agents and underpins their growing relevance in DeepTech systems.

The transition from classical to LLM-based agents introduces new architectural and governance challenges, particularly with respect to reliability, control, and long-term task execution in organizational environments. While LLM-based agents exhibit advanced reasoning and interaction capabilities, these features also raise questions about stability, accountability, and operational robustness when embedded in business workflows.

Analytical distinction between traditional agent architectures and contemporary LLM-based approaches is necessary. Section, therefore, examines the foundational characteristics of classical intelligent agents and contrasts them with LLM-based agents, establishing the conceptual basis for analyzing how agent architectures shape scalability, governance requirements, and business adoption of AI systems.

Russell and Norvig (2021) propose the Performance measure, Environment, Actuators, and Sensors (PEAS) framework as a structured method to specify AI agent task environments. Within PAES, classical AI agent architectures are commonly categorized by increasing complexity. Simple reflex agents rely solely on condition-action rules derived from the current perceptions and perform effectively only in fully observable and stationary environments. Model-based reflex agents enhance this logic by keeping an internal representation of the environment, allowing them to track unobservable states through transition and sensor models. While these architectures improve the agents' reliability, their behavior remains tightly constrained by predefined rules.

More advanced AI agents introduce goal-oriented reasoning. Goal-based agents incorporate explicit representations of the desired goals and employ search or planning algorithms to identify action sequences to achieve the specified goals. This approach offers greater flexibility than reflex-based architectures since agents can evaluate alternative paths toward an objective. At the same time, goal-based reasoning typically treats outcomes in binary terms, namely establishes whether a goal is achieved or not, limiting an agent's ability to address any potential trade-offs. Utility-based agents address this limitation by introducing functions that assign numerical values to states, enabling them to evaluate preferences and select actions that maximize expected utility under uncertainty (Russell & Norvig, 2021). This architecture is particularly relevant in environments where a balance between competing objectives, such as efficiency and safety, must be maintained.

In multi-agent contexts, the Belief-Desire-Intention (BDI) model has emerged as a prominent classical architecture. BDI systems explicitly represent an agent's beliefs about the world, its desires as goal states, and its intentions as committed plans of action (Meyer, 2014). By structuring their reasoning around practical decision-making, BDI agents exhibit autonomous and proactive behavior, enabling them to initiate actions rather than merely respond to external stimuli (Garro et al., 2025). These characteristics have made BDI architectures influential in early multi-agent system design.

Despite their conceptual clarity and strong theoretical control, classical agent architectures exhibit significant limitations when applied to contemporary business environments. Their reliance on symbolic logic, structured inputs, and predefined state spaces constrains their ability to operate effectively with unstructured data such as natural language, documents, or visual information (Tang et al., 2025). Moreover, table-driven and rule-based approaches become computationally infeasible as environmental complexity increases, owing to the exponential growth in the number of possible percept-action combinations (Russell & Norvig, 2021). From the perspective of this thesis, these limitations explain why classical agents struggle to scale beyond narrow, well-defined tasks and why they are ill-suited to operate as semi-autonomous actors within dynamic organizational workflows.

Technological progress in AI addresses these constraints by integrating LLMs into agent architectures. Compared to classical agents, which depend on explicit symbolic representations and predefined rules, LLM-based AI agents use probabilistic reasoning based on large-scale data. Standalone LLMs function primarily as stateless sequence-prediction systems, with their integration into agent architectures enabling them to serve as cognitive collaborators that interpret unstructured inputs, reason about context, and support goaldirected action in interactive environments (Tang et al., 2025). This shift is a decisive step toward agent systems that can function under uncertainty and variability, allowing stakeholders to use AI agents for controlled tasks and more complex, business-oriented goals.

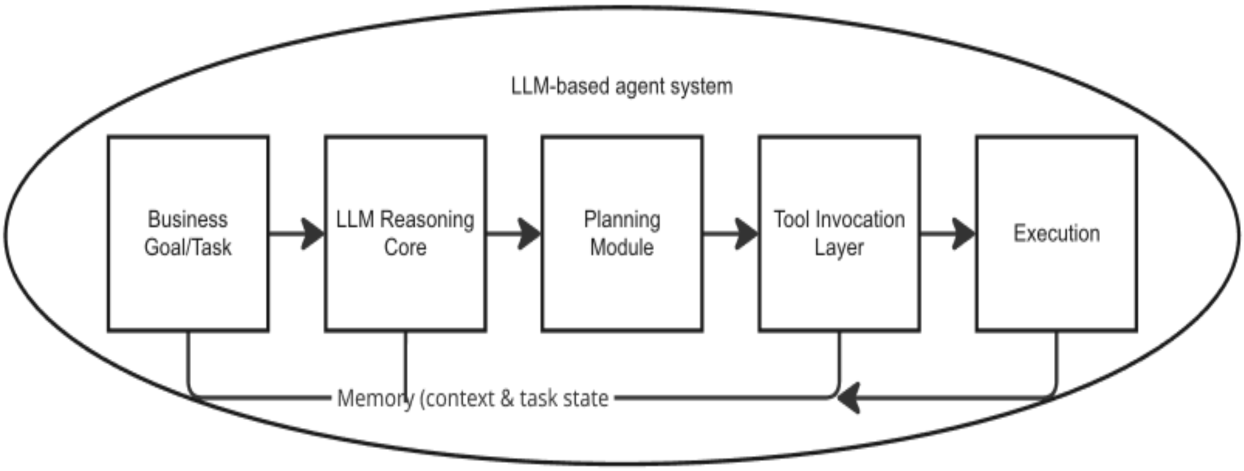
The key distinction between classical and LLM-based agents lies in their ability to process unstructured information and perform open-ended tasks without explicit rule encoding. LLM-based agents leverage general-purpose reasoning capabilities, allowing them to interpret ambiguous inputs, adapt to novel situations, and operate across diverse business contexts (Tupe & Thube, 2025). This flexibility enables agents to operate independently or collaboratively within complex systems, thereby extending their applicability beyond the narrow domains typically associated with traditional automation.

According to Tang et al. (2025), contemporary industry-grade LLM-based agents rely on three core architectural components that address the limitations of classical systems: memory, planning, and tool use. Memory mechanisms enable agents to maintain contextual continuity across extended task execution, overcoming the short-term coherence constraints of earlier architectures. These mechanisms range from immediate working memory to retrieval-based access to external knowledge sources, enabling agents to incorporate up-to-date, domain-specific information into their reasoning processes. Such memory augmentation supports learning from prior interactions and improves task performance over time.

Planning in LLM-based agents have fundamental differences from classical search-based approaches. Rather than exhaustively exploring predefined state spaces, they employ natural language reasoning to break-up high-level objectives into executable steps. Advanced planning techniques enable agents to generate, evaluate, and revise action sequences during execution, supporting greater adaptability in dynamic environments (Tang et al., 2025). This reactive planning capability allows them to have the flexibility to respond to unexpected conditions without complete reprogramming.

Tool use represents the final architectural pillar that distinguishes LLM-based agents from predictive models. By enabling agents to access external APIs, execute code, and interact with software systems, tool use extends agent capabilities beyond text generation toward direct operational control (Tupe & Thube, 2025). This functionality mitigates limitations associated with hallucination and outdated model knowledge by grounding agent actions in real-time data and system feedback. As a result, LLM-based agents transition from informational assistants to active participants in business processes. The internal coordination of these components within an LLM-based agent system is summarised in Figure 1.

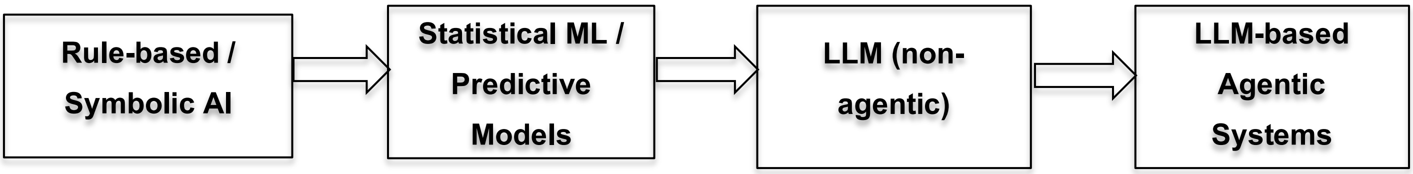
Figure 1. *Functional architecture of an LLM-based agent system*



Source: Compiled by the author

From an organizational perspective, these architectural developments enable the emergence of semi-autonomous systems capable of executing business tasks under human oversight. In contrast to classical agents, LLM-based architectures provide the technical foundation necessary for scaling agentic systems beyond experimental settings. This architectural progression of AI systems toward agentic architectures is summarized in Figure 2.

Figure 2. *Evolution of artificial intelligence systems toward agentic architectures*



Source: Compiled by the author

These architectural developments enable forms of semi-autonomous task execution that are a prerequisite for Level 2 AI adoption, in which agents are assigned responsibility for specific tasks while remaining subject to organizational governance and human oversight (Sohn, 2024). The following section builds on this distinction by examining how varying levels of agent autonomy translate into organizational adoption frameworks.

1.3. Taxonomies of Agent Autonomy and Adoption

The section integrates complementary taxonomies that link theoretical agent architectures with their practical adoption in DeepTech business environments. While previous sections examined how intelligent agents are technically designed, this section focuses on how increasing agent autonomy has allowed companies to adopt AI agents for commercial purposes.

Tang et al. (2025) provide a comprehensive maturity framework for LLM-based agents that categorizes systems by functional capabilities. The authors outline progress from basic process execution toward increasingly adaptive and socially interactive agent systems. At the center of this progression is the coordinated development of architectural components such as memory, planning, and tool use, which, combined, enable agents to operate with higher degrees of autonomy. This framework offers valuable insight into how technical capabilities evolve, but it focuses on the technology itself and does not fully address when such capabilities could become viable within business environments.

Sohn (2024) complements this perspective by introducing an organizational adoption framework that focuses on the redistribution of tasks between humans and AI systems. Without only focusing on technical maturity alone, this model has a step-by-step process, using which AI systems transition from supporting roles to independent task execution under human oversight. The framework uses organizational trust, control, and responsibility to highlight adoption thresholds that determine when agents can be used within routine business operations.

From the perspective of this thesis, neither framework is sufficient in isolation. Technical capability does not automatically translate into business adoption, while organizational readiness presupposes a certain level of agent functionality. This research, therefore, combines the maturity-oriented taxonomy of Tang et al. (2025) with the adoption-focused model of Sohn (2024) to identify the intersection at which AI agents can operate as semi-autonomous work companions. This integrated view provides the analytical basis for examining Level 2 AI adoption, where agents execute business tasks independently within defined governance boundaries.

Within this integrated perspective, the technological maturity of LLM-based agents is understood as a gradual progression rather than a binary shift from automation to autonomy. Tang et al. (2025) conceptualize this progression through a five-level capability maturity framework, which distinguishes increasingly complex forms of agent behavior and clarifies which levels of autonomy are technically viable for real-world organizational deployment.

At Level 1, agents function primarily as process execution systems. These agents translate natural-language inputs into formal instructions, such as code or database queries, and execute tasks through fixed, linear reasoning patterns, including Chain-of-Thought prompting (Tang et al., 2025; Zhao et al., 2025). Their operation relies on short-term working memory and assumes relatively stable environments. While effective for narrowly defined tasks, Level 1 systems remain fundamentally dependent on human direction and supervision.

Level 2 represents a shift toward interactive problem-solving systems, often described as copilots or advanced digital assistants. At this stage, agents can autonomously select tools and software interfaces and retrieve external information through retrieval-based memory mechanisms (Tang et al., 2025). Planning remains reactive rather than fully autonomous, as agents require human intervention to resolve errors, ambiguity, or unexpected outcomes (Yehudai et al., 2025). Although many commercially available systems are labeled as Level 2, their autonomy is typically constrained, and independent operation remains limited. From the perspective of this thesis, however, Level 2 constitutes the first point at which agents begin to execute business tasks with partial autonomy under human oversight.

Level 3 introduces end-to-end autonomous systems capable of independently performing complex domain-specific operations. These agents employ active-learning memory and global planning mechanisms that enable self-evaluation and correction during task execution (Tang et al., 2025). Addressing the “sim-to-real” gap is essential at this level, as agents must adapt to dynamic environments in which initial plans may fail (Zhao et al., 2025). While technically significant, Level 3 systems introduce heightened reliability and governance challenges that currently limit their widespread commercialization.

At Levels 4 and 5, agent autonomy extends beyond individual task execution to coordinated and socially adaptive systems. Level 4 focuses on multi-agent collaboration, where agents work together, exchange information, and coordinate workflows to manage complex tasks (Yehudai et al., 2025). Current evidence from operational experience suggests that this shift is not merely theoretical and by combining foundation models with multi-agent systems, Level 4 can enable greater autonomy in complex networks, while also creating new challenges related to coordination, reliability, and operational control that become noticeable beyond single tasks (Xu et al., 2024). Level 5 represents a theoretical endpoint in which agents generate their own objectives and evolve alongside human communities as adaptive socio-technical systems (Tang et al., 2025). Although elements of Levels 3 and 4 demonstrate technical feasibility, these stages remain largely experimental and fall outside the scope of stable business deployment.

In this thesis, the five-level framework is not studied as a linear roadmap toward full autonomy, but as a classification tool for identifying viable adoption scenarios. Despite the fact that Levels 3 and 4 showcase long-term technological potential, Level 2 is recognized as the critical inflection point for commercialization. At this level, agents begin to operate semi-autonomously within organizations, with humans still monitoring them. This distinction provides the technical foundation for the subsequent analysis of how organizations adopt and strategically scale AI agent frameworks.

While technical capability determines what AI systems can do, organizational adoption depends on how responsibilities, control, and accountability are distributed between humans and machines. Sohn (2024) argues that widely used classifications such as Artificial Narrow Intelligence (ANI), Artificial General Intelligence (AGI), and Artificial Superintelligence (ASI) are insufficient for business research because they focus on hypothetical cognitive benchmarks rather than on operational deployment. In response, Sohn proposes a Three Levels of AI Adoption framework that categorizes AI systems according to their functional role within organizational processes.

At Level 1: Augmentation, AI systems operate as supportive tools that enhance human performance without assuming responsibility for task execution. This stage aligns with the “copilot” model, in which AI assists users by providing recommendations, automating subtasks, or accelerating information processing, while humans retain full decision-making authority (Sohn, 2024). From a technical perspective, many Level 1 applications correspond to early LLM-based systems that generate value through assistance but lack autonomous execution capabilities (Tang et al., 2025). Although this level of AI agents can deliver immediate efficiency gains, it does not fundamentally alter organizational workflows or role structures.

Level 2: Collaboration and Equalization represent a qualitative shift in the logic of adoption. At this stage, AI systems assume responsibility for executing specific business tasks independently, functioning as a so-called “co-worker AI” or “agent AI” within defined operational boundaries (Sohn, 2024). Instead of merely supporting human activity, Level 2 AI agents take up tasks, interact with customers, validate processes, and conduct routine decision-making. Task execution is coordinated through architectural mechanisms that distribute work between humans, automated processes, and agents, enabling AI systems to operate semi-autonomously while still being subject to oversight. This level marks is where AI adoption requires organizations to address any governance, trust, and liability concerns before deploying such AI agents in real-world environments.

Level 3: Full Automation denotes a hypothetical stage in which AI systems assume control over all organizational operations, including the strategic coordination of human and automated resources. Sohn (2024) characterizes this level as “CEO AI,” reflecting a scenario in which AI directs organizational activity systematically. However, this stage mandates not only advanced technical capability but also social, legal, and institutional acceptance that exceeds the current technological and regulatory realities, which is why Level 3 remains mainly speculative, it might not be realistic that businesses would adopt such agents in the near term.

From the perspective of this thesis, the Three Levels of AI Adoption framework is not interpreted as a linear progression toward full automation. Instead, it serves as an analytical tool for identifying the organizational threshold at which AI systems transition from assistance to responsibility. Level 2 adoption is therefore the focal point of this research, as it represents the earliest stage at which AI agents begin to function as semi-autonomous actors within business processes. This distinction is critical for understanding the strategic, organizational, and governance challenges associated with scaling AI agent frameworks in DeepTech ventures.

This research conceptualizes Level 2 Organizational Adoption, referred to as AI Equalization, as a central challenge in commercialization for DeepTech entrepreneurs. While organizational adoption frameworks describe how AI systems are embedded into business roles, the preceding analysis demonstrates that such adoption is contingent upon achieving a sufficient level of underlying technical maturity. By synthesizing the technical capability taxonomy proposed by Tang et al. (2025) with the organizational adoption model developed by Sohn (2024), this section clarifies the conditions under which AI agents can transition from supportive tools to semi-autonomous organizational actors.

Comparing these two frameworks shows that Level 2 organizational adoption cannot be achieved through Level 2 technical capability alone. Even though Tang et al. (2025) classify Level 2 agents as interactive problem-solving systems, these agents remain primarily reactive and dependent on human intervention when encountering uncertainty, system errors, or novel situations. In contrast, organizational Level 2 adoption assumes that AI systems can take responsibility for executing business tasks. These can extend beyond routine assignment completion and require AI agents to manage exceptional cases, maintain contextual continuity, and interact reliably with multiple systems within a business. As a result, AI Equalization requires businesses to have technical capabilities that extend beyond the reactive boundaries of Level 2.

From a technical perspective, this implies that organizations must incorporate elements associated with higher levels of agent maturity. Persistent and active-learning memory mechanisms, associated with Level 3 systems, are necessary for keeping institutional knowledge, which enable agents to learn from past interactions and not solely rely on short-term memory windows (Tang et al., 2025). Similarly, planning and sufficient feedback capabilities become essential since agents operating as organizational co-workers must be able to detect failures, revise plans, and resolve breakdowns without constant human intervention (Sohn, 2024; Zhao et al., 2025). Without these capabilities, AI agents remain limited to assistive functions and cannot sustain independent role execution.

Furthermore, many business roles involve coordination across processes, systems, and stakeholders whose demands exceed the capacity of a single agent. In such cases, the reliable execution of organizational tasks depends on multi-agent collaboration and orchestration mechanisms typically associated with Level 4 technical maturity (Tang et al., 2025). These mechanisms allow agents to distribute workloads, exchange information, and coordinate actions, thereby supporting the inherently complex real-world business operations (Yehudai et al., 2025). Consequently, while organizational Level 2 adoption does not require fully autonomous socio-technical systems, those AI agents often demonstrate technical capabilities that are exhibited at Levels 3 and 4.

This synthesis directly leads to a key insight that underpins the analytical focus of this thesis, namely that the scaling of AI agent frameworks does not mean linear progress through technical autonomy levels. Instead, scaling becomes a challenge for enterprises as they try to bridge the gap between technically reactive agents and technically autonomous or collaborative systems to enable organizational AI Equalization. Knowing this, commercialization relies on aligning advanced technical capabilities with organizational structures that allow agents to operate independently while remaining accountable within the boundaries of an enterprise's management structure.

The relationship between technical capability and organizational adoption is summarized in Table 1, which compares the two frameworks across key dimensions, including unit of analysis, source of autonomy, human role, and limiting factors. This comparison reinforces the argument that AI Equalization emerges at the intersection of technological maturity and organizational governance, and not a direct consequence of progress in either domain alone. The shift from predictive assistance to semi-autonomous operation, therefore, defines the current DeepTech entrepreneurship context examined in the following section.

Table 1. *Comparison of technical capability and organizational adoption frameworks*

Dimension	**Technical Capability (Tang et al.)**	**Organizational Adoption (Sohn)**
Unit of analysis	AI system architecture	Business role allocation
Role of AI	Autonomous execution engine	Assigned organizational actor
Source of autonomy	Technical system	Organizational governance
Human role	Supervisor of execution	Accountability holder
Limiting factor	Model reliability and planning	Trust, liability, and control

Source: compiled by the author, based on Tang et al., 2025, and Sohn, 2024

1.4. The DeepTech Entrepreneurship Context

Having established the technical and organizational conditions under which AI agents can operate at Level 2 adoption, the analysis now places these conditions within the broader context of DeepTech startups. Scaling AI agent frameworks toward Level 2 adoption introduces challenges that are unlike those encountered in conventional software development. Unlike typical digital products, AI agents that are operationally independent within business processes require not only functional scalability but also reliability, governance, and infrastructural robustness. Research highlights that many AI technologies fail to turn theoretical performance into sustained commercial value due to the complexity and organizational constraints related to the integration of AI agents (Zhan et al., 2024).

DeepTech ventures operate in environments that could be described as highly capital-intensive with long development cycles and strong dependence on advanced infrastructure such as cloud computing and foundation models (Luitse, 2024). These features can create structural barriers to scaling, particularly for startups that lack the resources to develop proprietary infrastructure. At the same time, talent scarcity and the need for knowledge across multiple disciplines further complicate the commercialization of autonomous systems, while reliability concerns could become problematic for AI agents since errors or unpredictable behavior can directly affect operational outcomes and organizational trust (Sohn, 2024).

This thesis focuses on the DeepTech entrepreneurship context as the intersection of advanced technical requirements and constrained strategic flexibility. Understanding these conditions is essential for understanding how AI agent ventures navigate the transition from experimental systems to them becoming commercially viable. These challenges are not unique to AI agents but reflect the broader state of DeepTech ventures, where technological uncertainty, infrastructure dependence, and governance constraints shape the path to commercial scale.

DeepTech ventures differ fundamentally from conventional digital startups in the risks they face and the resources they require. Whereas typical digital startups primarily confront market risk - uncertainty related to customer demand, pricing, or distribution - DeepTech ventures are characterized by high levels of technological risk arising from unresolved scientific and engineering challenges (Adekunle et al., 2024). As a result, such ventures often experience extended research and development cycles before reaching a level of technical maturity suitable for commercialization.

This distinction is particularly relevant in the context of AI agent development. Entry-level AI applications, such as basic chatbots or decision-support tools corresponding to Level 1 adoption, can often be deployed using existing APIs and pre-trained models with relatively low technical barriers. In contrast, the development of industry-grade AI agents capable of memory persistence, planning, and tool use - corresponding to Levels 2 and 3 technical maturity - introduces significantly greater complexity (Tang et al., 2025). These systems must operate reliably in dynamic, real-world environments, requiring solutions to challenges such as the “sim-to-real” gap and the coordination of complex agent interactions.

Consequently, ventures developing AI agent frameworks are exposed to the defining characteristics of DeepTech entrepreneurship, including substantial capital requirements, dependence on advanced infrastructure, and constraints on rapid iteration. Unlike software startups that can pivot quickly in response to user feedback, DeepTech ventures must often resolve foundational technical uncertainties before meaningful market validation is possible (Adekunle et al., 2024). From the perspective of this thesis, this technological uncertainty explains why scaling AI agent frameworks toward Level 2 adoption constitutes not only a technical challenge but also an entrepreneurial one shaped by long development horizons and limited strategic flexibility.

Within the AI agent domain, these entrepreneurial constraints manifest as structural characteristics that shape the transition from experimental models to scalable commercial systems. Beyond market uncertainty, ventures face challenges stemming from capital intensity, talent scarcity, and the reliability requirements inherent in real-world deployment. A macroeconomic perspective reinforces this view by emphasizing that the economic impact of AI depends on adoption conditions and complementary investments, rather than following automatically from improvements in model capability alone (Acemoglu, 2024).

One of the main features of AI agent ventures is their dependence on capital-intensive computational infrastructure. Unlike conventional software startups, which can often scale using relatively modest computing resources, AI agent systems rely on LLMs that require specialized hardware, such as graphics processing units (GPUs) or tensor processing units (TPUs), as well as large-scale data storage and processing capabilities (Luitse, 2024). As a result, DeepTech entrepreneurs typically rely on hyperscale cloud providers, such as Amazon Web Services, Microsoft Azure, or Google Cloud, to access the infrastructure required to train and operate AI agentic systems. This dependency creates asymmetries in bargaining power, with platform owners retaining control over access conditions, pricing models, and technical standards that directly affect the strategic options available to startups (Canboy & Khlif, 2025). Infrastructural dependence becomes a structural constraint on scaling.

In addition to infrastructure constraints, AI agent ventures face pronounced challenges related to talent scarcity and organizational readiness. The development and deployment of agentic systems require interdisciplinary expertise that spans machine learning, data engineering, system architecture, and domain-specific knowledge. Organizations frequently struggle to attract and retain such talent, which limits their capacity to scale development and support enterprise-grade deployments (McKinsey, 2025). Beyond technical skills, successful adoption of AI agents also necessitates organizational adaptation, including the redesign of workflows, governance structures, and data management practices. As Sjödin et al. (2023) note, firms seeking to commercialize advanced AI solutions must often address internal capability gaps alongside those solutions engaging with external stakeholders, particularly when serving enterprise clients with complex operational requirements.

Another critical constraint in the AI agent domain is the constant presence of the "sim-to-real" reliability gap. Research indicates that agents that perform well in controlled or simulated environments often fail when exposed to the variability of real-world organizational settings (Tang et al., 2025). This limitation stems from the interdependence among architectural components, including memory, planning, and tool use, whereby deficiencies in any single component can lead to a systemic task failure. Tang et al. (2025) describe this vulnerability using the "wooden barrel" analogy, namely that a system's performance is constrained by its weakest element. For ventures considering Level 2 adoption, these reliability concerns could be a decisive barrier, with operational environments requiring consistent and predictable system behavior.

Taken together, these characteristics illustrate why AI agent ventures exemplify challenges faced by DeepTech entrepreneurs. Capital intensity, infrastructural dependence, talent constraints, and reliability risks all play a part in slowing down commercialization and increasing uncertainty for such startups. These conditions help explain why scaling AI agent frameworks toward Level 2 adoption is not only a matter of technological improvement, but a complex organizational process shaped by structural limitations. These same constraints directly shape the commercialization phase, in which ventures must transform experimental systems into stable, deployable solutions that would be used in real-world business environments.

In practice, the structural characteristics of DeepTech ventures translate into significant commercialization challenges when startups attempt to convert experimental AI agent frameworks into reliable enterprise-level systems. This transition is commonly described as the "Valley of Death," referring to the period between achieving a successful proof of concept and delivering a stable, market-ready product capable of sustained operation. For AI agent ventures targeting Level 2 adoption, this phase is particularly critical, as it requires agents to move from demonstrative performance to dependable role execution within organizational processes.

One of the key obstacles in this transition is establishing organizational trust. To function as co-workers at Level 2 adoption, AI agents must be seen as reliable, predictable, and governable. However, deep learning models that are primarily agentic systems are often characterized as black boxes, as their internal decision-making processes remain largely opaque (Black et al., 2024). This lack of transparency complicates accountability and affects managerial confidence in delegating operational responsibilities to AI systems. In addition to interpretability concerns, agentic systems are exposed to two broad categories of risk:

technological failures and system-level design vulnerabilities. These risks can undermine human oversight and lead to unintended interactions, particularly in sensitive domains such as financial services or defense-related applications (Kilian, 2025). Reliability concerns can also be interpreted through a resilience perspective. When agents are embedded in interconnected service networks, the stability of the broader system becomes an explicit operational requirement rather than a property of an isolated model or pilot workflow (Geng & Liu, 2025). Importantly, this implies that commercialization depends on evidence-making practices that demonstrate performance under realistic conditions and clarify acceptable thresholds of reliability, rather than relying on anecdotal pilot success; validation-oriented studies combining large language models with agent-based systems illustrate why systematic evaluation becomes a prerequisite for credible deployment (Peasley et al., 2025).

Addressing these risks imposes substantial cost and complexity on the commercialization process. Entrepreneurs are required to invest in safety mechanisms, operational guardrails, and compliance systems that constrain agent behavior and mitigate issues such as hallucinations or unauthorized actions (Google Cloud, 2025). While these measures are essential for the deployment of AI systems within businesses, they prolong development timelines and raise the threshold for market entry, disproportionately affecting resource-constrained DeepTech startups.

Beyond technical reliability, commercialization could be further slowed by organizational and ecosystem-level coordination challenges. Although many firms initiate pilot projects or proof-of-concept (POC) initiatives, only a limited number succeed in deploying AI agents at scale. Effective implementation requires robust technical infrastructure and the alignment of motivation across a network of stakeholders, including customers, partners, and platform providers (Sjödin et al., 2023). This alignment is challenging to achieve for AI agent-based systems, as their value often depends on integration across multiple organizational networks.

From a strategic perspective, the main challenge of commercialization is to enable agents to perform reliably at scale while minimizing the need for continuous human supervision. Achieving this balance requires businesses to develop strong capabilities for both value discovery and value realization, connecting advanced technical functionalities with concrete business needs (Sjödin et al., 2023). The need to resolve reliability issues alongside organizational governance and control mechanisms helps explain why many firms remain at Level 1 adoption, despite the availability of more advanced technologies. As Sohn (2024) argues, Level 2 AI adoption represents a qualitative shift rather than an incremental improvement, with which comes the need for organizations to redefine responsibility, trust, and control in the relationship between humans and AI.

In this thesis, these commercialization challenges are interpreted as defining features of the DeepTech AI entrepreneurship context. The difficulty of crossing the Valley of Death in agent-based systems underscores why Level 2 adoption remains rare and why scaling AI agent frameworks constitutes a distinct entrepreneurial problem. Understanding these challenges provides the foundation for the subsequent analysis of business models, platform dependence, and strategic responses employed by AI agent ventures.

1.5. Business Models and Platform Ecosystems

The successful commercialization of AI agent frameworks requires DeepTech ventures to navigate the intersection of business model design and digital platform ecosystems. While the previous sections established the technical maturity needed for Level 2 AI adoption and the structural constraints of DeepTech entrepreneurship, commercialization ultimately depends on how these technologies transition into valuable propositions and sustainable revenue mechanisms (Teece, 2010).

Unlike traditional software products, AI agent frameworks derive value from continuous interaction, data processing, and deep integration into business processes. As a result, their commercialization is inherently tied to platform-based environments that control access to users, data, computational infrastructure, and complementary services. Digital platform ecosystems are known for their multi-sided interactions, network effects, and governance mechanisms that shape how value is created, distributed, and controlled among participants (Parker et al., 2016; Hein et al., 2020). For AI agent ventures, engagement with such ecosystems is rarely optional, as access to foundation models, cloud infrastructure, and distribution channels is typically controlled by platform owners.

The following section examines how DeepTech ventures design and adapt business models for AI agent frameworks within platform-dominated environments. It analyzes how platform policies influence the independence of enterprises by shaping value creation, value capture, and control over key resources. By linking business model theory with the governance of the ecosystem of platforms, this chapter establishes the conceptual understanding of the strategic trade-offs faced by AI agent ventures as they seek to scale toward Level 2 adoption.

A business model defines how a firm creates, delivers, and captures value by linking the value proposition to customer relationships, channels, revenue streams, and the underlying cost structure. Osterwalder and Pigneur (2010) conceptualize a business model as an integrated system of interdependent elements, providing a suitable analytical perspective to examine how AI-driven ventures transform technological capabilities into coherent value systems. In the AI business, this perspective is increasingly applied to argue that ventures ’ competitive advantage depends not only on model performance but also on how AI capabilities are embedded into repeatable mechanisms of value creation and value capture (Jobstreibizer et al., 2025).

The AI sector has accelerated the shift from one-time product sales toward continuously operating service-delivery systems. This transition is enabled by AI becoming a part of the processes of servitization, whereby value is generated through ongoing customer interaction and information exchange rather than individual transactions (Canboy & Khlif, 2025). Within this context, Sjödin et al. (2023) make the division between augmentation models, which enhance human decision-making, and automation models, which delegate task execution to autonomous systems to improve operational performance. While both help explain how AI-based ventures achieve efficiency gains, improved decision quality, and scalable personalization (Irman et al., 2025), they also suggest that value realization depends on the organizational and commercial mechanisms through which these benefits are delivered consistently and at acceptable cost for businesses. As such, extracting value from AI solutions requires more than technical capability alone, mandating a business model aligned with the logic of profiting from innovation, enabling innovators, and not imitators or dominant asset holders, to claim any capital returns (Teece, 2010).

For AI agent frameworks, this challenge is further shaped by data-driven network effects. As agent performance and utility improve through accumulated interaction data, continued user participation becomes a practical condition for both learning dynamics and value capture (Parker et al., 2016). Sustained commercialization, therefore, depends not only on initial adoption but also on acquisition and retention dynamics that determine recurring system use over time. Agent-based approaches have been used to model these dynamics in startup contexts, underscoring the importance of treating user participation as a system-level variable of commercialization rather than a one-time transaction outcome (Sayyed-Alikhani et al., 2021).

At the same time, AI agent ventures rarely operate as independent market actors. Instead, they are increasingly embedded within digital platform ecosystems that mediate access to infrastructure, users, and complementary services. A digital platform ecosystem consists of a platform owner that governs interactions with an ecosystem of autonomous complementors and consumers through formal and informal governance mechanisms (Hein et al., 2020). These ecosystems are shaped by network effects, whereby the value of participation increases as additional users and producers join the platform (Parker et al., 2016). For AI agent ventures, such effects can accelerate diffusion and adoption. Still, they also reinforce structural dependence on platform-controlled environments, directly conditioning how business models can be designed and scaled.

Access to platform ecosystems is enabled through so-called boundary resources, such as Application Programming Interfaces (APIs) and Software Development Kits (SDKs), which structure the arm’s-length relationship between platform owners and external developers (Ghazawneh & Henfridsson, 2013). These boundary resources are designed to promote generativity by allowing third-party actors to develop new applications and services without direct involvement from the platform owner (Hein et al., 2020). In the context of AI agents, generativity is particularly important, as agent frameworks often require deep integration with data sources, enterprise software, and external tools to deliver value.

At the same time, platform ecosystems must preserve system integrity by ensuring stability, security, and control over core functionalities. This creates a persistent tension between enabling external innovation and maintaining platform reliability, which platform owners manage through governance mechanisms that regulate access, usage rights, and technical standards (Engert et al., 2025). For AI agent ventures, this trade-off is especially consequential, as restrictions imposed to protect platform integrity can limit agent autonomy, constrain scalability, or introduce uncertainty regarding future access conditions. This integrity requirement becomes even more salient when agentic systems rely on distributed or multi-agent coordination, where agents exchange information across organizational or technical boundaries; in such settings, privacy and security constraints can shape the feasibility of coordination itself, motivating privacy-preserving approaches to consensus and distributed optimization in multi-agent systems (Wang et al., 2023).

Empirical research suggests that small and medium-sized enterprises (SMEs) often choose to participate in existing platform ecosystems because they lack the resources required to develop independent digital infrastructures for AI deployment (Masiero et al., 2024). This reliance is consistent with evidence that business-to-business (B2B) SMEs frequently use AI platforms as a practical pathway to integrate AI technologies that would otherwise demand substantial internal capabilities and investment (Wei & Pardo, 2022). While participation in external platforms can lower entry barriers and accelerate innovation, it simultaneously puts entrepreneurial activity within governance structures controlled mainly by the owners of such platforms. Yu and Sekiguchi (2024) describe this condition as “platform-dependent entrepreneurship,” highlighting how ventures rely on platforms for distribution and core operational capabilities.

These platform ecosystem dynamics represent a double-edged sword for AI agent ventures. While platforms enable access to critical infrastructure, data, and markets necessary for scaling AI agent frameworks, at the same time, they constrain strategic autonomy through governance mechanisms and dependency structures. As Yu and Sekiguchi (2024) suggest, in platform-dependent enterprises, their ability to operate and scale becomes contingent on continued access to platform-controlled resources and interfaces. When this happens, platform owners retain disproportionate authority over access conditions, technical standards, and rules of participation, creating dependencies that may become a strategic vulnerability as ventures scale.

These dynamics can affect AI agent frameworks, where agents increasingly mediate core organizational tasks rather than peripheral functions. As agentic systems assume more responsibility for the execution, coordination, and decision-support of tasks across workflows, platform governance choices begin to shape the technological access and permissible scope of delegation, accountability, and control within organizations using AI agents. Thus, the scaling environment can be interpreted as an emerging agent economy, in which AI agents function as operational actors and where institutional rules, incentives, and governance arrangements determine how agent-based value is produced and appropriated (Hadfield & Koh, 2025). Platform participation becomes not only a technical or distributional choice, but a structural condition that directly influences business model frameworks, exposure to risks, and the feasibility of scaling AI agents toward Level 2 adoption.

This concentration of power is particularly evident in the AI domain due to the centralized nature of AI infrastructures. Canboy and Khlif (2025) argue that AI platform providers exert control through the ownership of data distribution mechanisms and the algorithmic standards that structure ecosystem participation. Platform owners that control foundational infrastructure components, whether it would be cloud computing or foundation models, can impose technical requirements, pricing schemes, and usage constraints that developers have no choice but to accept to participate (Luitse, 2024). These dynamics often result into vendor lock-in and reduced strategic autonomy, shaping how ventures build their systems and which business models remain feasible under platform-imposed constraints. In an agent-based economy, these things matter because in practice, the platform’s rules constitute a form of institutional infrastructure since they determine what kinds of agentic interaction are possible, how work can be delegated, and where value is captured (Hadfield & Koh, 2025).

For companies seeking to scale AI agent frameworks toward Level 2 adoption, platform dependence therefore creates a strategic dilemma. While access to hyperscaler platforms is often necessary to achieve the computational scale, reliability, and deployment tooling required for AI agentic systems (Luitse, 2024), deeper integration with a single platform increases risks related to unilateral policy changes, pricing shifts, and technical restrictions that can undermine the long-term viability of AI agents. Foster (2024) adds nuance to this dependency by framing “openness” as a technical attribute as well as a downstream governance condition that shapes the extent of autonomy complementors retain when using AI agents. In practice, even when platforms support forms of openness, the rules and interfaces that shape access can still constrain downstream innovation and value capture, conditioning the participation in the ecosystem on the policies set upstream (Foster, 2024).

Research identifies several strategies through which platform-dependent entrepreneurs attempt to mitigate these risks. Yu and Sekiguchi (2024) highlight multi-homing as a common approach used by enterprises, whereby ventures deploy services across multiple platforms to reduce dependency on any single provider and introduce redundancy. Additional strategies include developing proprietary intellectual property and forming collaborative arrangements with other ecosystem actors to strengthen bargaining positions, minimizing vulnerability to unilateral platform decisions. However, the feasibility of these strategies is constrained by resource limitations and technical complexity, particularly when it comes to DeepTech startups operating under capital and talent constraints. At such companies, parallel deployments and portability engineering can create non-trivial costs that are difficult to absorb when capital is limited.

Therefore, managing platform dependence requires developing specific dynamic capabilities. Warner and Wäger (2019) emphasize capabilities related to detecting platform changes and adapting internal structures accordingly, while Teece (2007) highlights the strategic value of maintaining flexibility in environments characterized by rapid technological and market shifts. In the context of AI agent ventures, these capabilities enable the monitoring of any potential platform governance changes, redesigning integration architectures, and the preservation of continuity of value creation despite shifting platform rules of access and control.

This thesis looks at governance and power dynamics within platform ecosystems as a critical part of the entrepreneurial environment in which AI agent frameworks are scaled. Platform architecture and governance interact directly with entrepreneurial decision-making, shaping how ventures design business models, allocate resources, and pursue Level 2 AI adoption. These interactions influence the strategic landscape examined in the empirical part of the research, where the experiences of ecosystem stakeholders, who navigate platform dependence, are analyzed in detail.

1.6. Actor Roles in AI Platform Ecosystems

AI agent frameworks can enter the market as technological products, yet their commercialization depends on how they are used within organizational and platform-based systems that shape stakeholder interaction and responsibility. The theory of platform ecosystems emphasizes that value in digital environments is produced through the coordinated activity of heterogeneous stakeholders being at varying distances from the platform’s core (Ghazawneh & Henfridsson, 2013). Stakeholders have different capabilities, functions, and interdependencies, which together influence innovation trajectories, adoption pathways, and ultimately the scalability of agent-based solutions.

Within AI platform ecosystems, three analytically distinct actor roles can be identified: creators, integrators, and adopters. Creators develop foundational AI technologies, including foundation models, agent architectures, and underlying computational infrastructure. By defining core system architectures and interfaces, these actors establish the technological boundaries within which subsequent innovation occurs (Hein et al., 2020). Their control over essential resources positions them as central actors in shaping ecosystem evolution and determining which forms of agent development become technically feasible and economically viable.

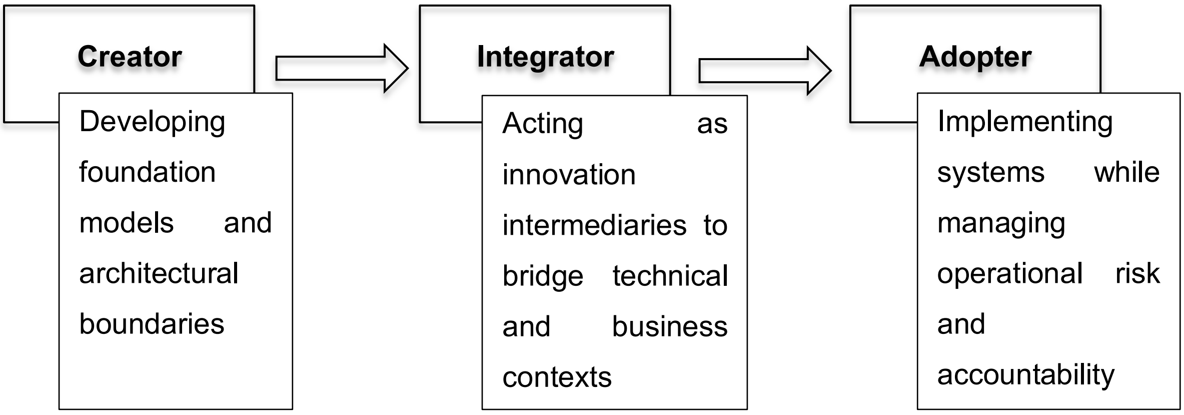
Integrators operate at the interface between core AI technologies and application-specific contexts. Their role involves combining AI components with domain knowledge, organizational processes, and existing information systems to produce deployable solutions (Ghazawneh & Henfridsson, 2013). This function becomes particularly critical when general-purpose AI models are translated into reliable agentic systems capable of operating within complex business environments. Significantly, the integrator role extends beyond implementation in a narrow technical sense. Innovation research conceptualizes such actors as innovation intermediaries whose contribution lies in reducing coordination and knowledge barriers across organizations and ecosystems. Caloffi et al. (2023) emphasize that intermediaries perform heterogeneous functions rather than a single service task, supporting adoption by connecting actors, structuring interaction, and facilitating the transfer of technologies into practical use. Complementing this view, Colović et al. (2024) highlight that innovation intermediaries become especially important in contexts shaped by emerging digital technologies, where uncertainty, rapid tool evolution, and uneven organizational readiness increase the need for translation, validation, and orchestration across stakeholders. In the context of AI agent frameworks, these perspectives imply that integrators shape not only deployment choices but also how agent capabilities are operationalized, governed, and sustained across heterogeneous infrastructures.

Adopters are organizations that implement AI agent systems within their operational environments. Their responsibilities extend beyond technical deployment to include governance, accountability, and performance management. Adopters must assess the risks associated with delegating tasks to AI agents and ensure that systems align with organizational objectives, regulatory requirements, and internal policies. As a result, adopters are directly exposed to the consequences of system failures, misalignment, or improper use, making their perspective central for understanding real-world adoption outcomes and the practical limits of autonomy.

Differentiating between these actor roles is analytically significant rather than merely descriptive. Prior research demonstrates that power asymmetries, governance mechanisms, and innovation incentives differ across ecosystem layers (Hein et al., 2020). While creators exert influence through architectural control and infrastructural ownership, integrators face challenges related to coordination and dependency management across platforms and organizations, and adopters bear operational, compliance, and reputational risks associated with deployment of AI agents. These differences shape how actors evaluate readiness for adoption, determine acceptable autonomy thresholds, and respond to scaling constraints.

In this thesis, the categorization of creators, integrators, and adopters functions as a theoretical framework for the empirical analysis. Examining AI agent scaling through role-specific lenses enables the research to capture how technical capabilities, business constraints, and platform governance interact across the ecosystem. This perspective is especially relevant for analyzing the transition from Level 1 to Level 2 AI adoption, where agents begin to execute semi-autonomous business functions under managerial supervision. Understanding how each actor role navigates this transition provides the conceptual basis for the empirical investigation presented in the following chapters. The functional interdependencies and a flow between technology creators, integrators, and adopters, which characterize the progression from foundational development to organizational deployment, are illustrated in the AI agent ecosystem actor interactions Figure 3.

Figure 3. *AI agent ecosystem actor interactions*


Creator
Developing foundation models and architectural boundaries
Integrator
Acting as innovation intermediaries to bridge technical and business contexts
Adopter
Implementing systems while managing operational risk and accountability

Source: Compiled by the author

1.7. Dynamic Capabilities for Scaling AI Frameworks

Transforming experimental AI models into commercially viable agent frameworks requires more than technical expertise. The transformation demands strategic capabilities that enable organizations to adapt when facing technological uncertainty and rapid change in business environments. While operational capabilities can allow firms to execute established activities efficiently, dynamic capabilities enable the renewal of competencies, the reconfiguration of resources, and timely responses to emerging opportunities (Teece, 2007). In AI agent contexts, this renewal frequently entails changes in how value is created, delivered, and captured, suggesting that any changes to a business model should be treated as an integral process of dynamic capability development rather than an aftereffect of technical progress (Juntunen, 2017). These capabilities are particularly important for DeepTech ventures which operate in fast-evolving and platform-dependent ecosystems (Warner & Wäger, 2019).

Scaling AI agent frameworks toward Level 2 adoption, or AI equalization, requires organizations to move beyond experimentation of the agents and establish structures that support semi-autonomous role execution under supervision (Sohn, 2024). This transition depends on the ability to identify viable opportunity spaces, commit resources to credible commercialization pathways, and reconfigure internal structures so that agentic systems can operate reliably within organizational and ecosystem constraints (Teece, 2007; Warner & Wäger, 2019). The following subsections apply the sensing-seizing-transforming logic to specify how such scaling is managed in practice.

First among the dynamic capabilities, sensing refers to the ability to identify emerging opportunities and threats before they fully materialize (Teece, 2007). For DeepTech ventures developing AI agent frameworks, sensing is intensified by the pace of change in foundation models, tooling ecosystems, and platform governance arrangements. In digital contexts, this requires digital scouting and scenario planning that track technological developments and ecosystem stability, enabling businesses to anticipate shifts that affect integration feasibility and scalability (Warner & Wäger, 2019). For AI agents, this includes monitoring developments in model capabilities, changes in access conditions, and the maturity of complementary tools required for the execution of tasks.

Sensing also requires value-discovery capabilities that connect technological possibilities to concrete operational problems and to measurable performance improvements (Sjödin et al., 2023). From the perspective of this thesis, a recurring strategic failure in emerging AI domains is to treat “automation feasibility” as equivalent to “commercial desirability.” For agentic systems, this distinction is crucial because Level 2 adoption implies costs and risks that extend beyond model performance, including supervision load, integration overhead, and governance requirements (Sohn, 2024). To avoid technology-driven experimentation without a commercialization pathway, sensing therefore benefits from explicit evaluation logic. Kemell et al. (2020) emphasize disciplined validation under uncertainty in early-stage technology development, while Ziakis and Kavoura (2025) reinforce that initiatives require credible evaluation perspectives to support decisions about feasibility and value. In this context, sensing should include the capacity to define decision criteria that reflect not only functional performance but also operational sustainability and risk exposure.

Finally, sensing in agent ventures must include early identification of structural risks. These include platform policy changes, shifts in dependency exposure, and the continuous presence of the sim-to-real gap, in which agent performance observed in controlled settings fails to be reliable in real-world scenarios (Teece, 2007; Tang et al., 2025). Therefore, sensing functions as the selection mechanism that determines whether scaling efforts concentrate on commercially meaningful workflows or remain locked in impressive but fragile prototypes.

The second capability is seizing, namely capturing opportunities, mobilizing resources, and committing to courses of action that enable value capture (Teece, 2007). When examining AI agent frameworks, seizing translates technological potential into commercially viable solutions under uncertainty, platform dependence, and evolving organizational requirements. Strategic agility is crucial since rapid prototyping and iterative testing support validation of viable workflows can expose operational constraints that were not observed during laboratory demonstrations (Warner & Wäger, 2019). For agentic systems, seizing is inseparable from structured validation, because scaling decisions depend on whether reliability, integration feasibility, and governance constraints can be managed consistently when implemented into organizational workflows.

This is where evaluation becomes commercially decisive rather than being just a methodological “nice-to-have.” Kemell et al. (2020) strengthen the argument that progress under uncertainty depends on explicit validation and not assumed learning, and Ziakis and Kavoura (2025) similarly emphasize the need for such evaluation approaches that enable credible assessment of outcomes and value. Seizing in agent ventures includes the ability to establish evidence-based thresholds, such as acceptable reliability, supervision effort, and integration robustness, which make the decision not to use or to use an AI agent and scaling decisions defensible in front of stakeholders.

Seizing also depends on collaboration across value chains. Hafeez et al. (2025) argue that complex digital commercialization often requires co-creation with customers and partners to identify integration points, data access arrangements, and operational responsibilities. For AI agents, such collaboration is not merely supportive but structurally necessary, because role-like execution requires clarity on accountability boundaries and escalation logic within existing business processes.

From a business model perspective, seizing requires value-realization capabilities that align incentives and structure offerings, combining automation and augmentation (Sjödin et al., 2023). As agents approach Level 2 adoption, organizations must design arrangements that redistribute work between humans and agents while addressing trust, accountability, and changes in employee roles (Sohn, 2024). Juntunen (2017) supports this by framing business model change as a dynamic capability: in agentic contexts, seizing entails the capacity to redesign value-delivery and capture mechanisms so that supervision costs, accountability structures, and role reallocation are reflected in what is promised, priced, and governed.

At the same time, seizing decisions are shaped by appropriation risks and platform policies. Holland et al. (2024) state that firms must define their business scope to maximize the value of innovation while limiting exposure to external stakeholders controlling critical infrastructure and channels. For AI agent ventures, this makes seizing include strategic boundary decisions about what to develop internally, what to source externally, and what must be owned to protect long-term value capture.

Lastly, transformation (reconfiguration) enables firms to sustain innovation by reshaping internal structures and relationships with the ecosystem over time (Teece, 2007). For AI agent frameworks, transforming becomes critical after initial commercialization because scaling requires repeatable adoption rather than isolated pilot trials. At the ecosystem level, Hafeez et al. (2025) describe ecosystem revamping capabilities through which firms restructure partner networks to access complementary technologies, expertise, and channels. In AI agent ventures, such restructuring supports expansion and can also reduce risks created by narrow dependencies.

Internally, transforming requires redesigning processes, governance arrangements, and decision-making structures to accommodate semi-autonomous systems (Warner & Wäger, 2019). Continuous value optimization is crucial because real-world deployments expose limitations not evident during early trials and could require systematic refinement of reliability and customer fit (Sjödin et al., 2023). When looking at this issue in the thesis, transformation is where ventures either institutionalize AI agentic workflows into durable operations or remain trapped in repeated pilots that do not materialize into concrete solutions due to weak governance practices and insufficient operational knowledge.

Finally, to bridge the gap between technical maturity and organizational adoption, dynamic capabilities are essential. Tang et al. (2025) show that Level 3 technical maturity can include end-to-end planning and tool-use capabilities, yet these features alone do not guarantee business value or organizational integration. From a managerial perspective, Level 2 adoption - AI equalization - represents the point at which AI systems assume responsibility for defined business roles under supervision (Sohn, 2024). Achieving this shift requires the coordinated deployment of sensing (selecting viable roles and constraints), seizing (mobilizing resources and designing commercialization logic), and transforming (reconfiguring governance and workflows to sustain hybrid human-agent operations) (Teece, 2007; Warner & Wäger, 2019).

A critical implication emphasized by this thesis is that Level 2 AI agent adoption is not built on technical capability claims but on evidence-based operationalization. Kemell et al. (2020) and Ziakis and Kavoura (2025) support the position that scaling decisions under uncertainty require credible evaluation and validation routines. In agent contexts, this means defining concrete thresholds for reliability and supervision effort and aligning business model design with the realities of task reallocation, accountability, and governance constraints (Sjödin et al., 2023). A conceptual business model can change as dynamic capabilities clarify why this alignment is necessary. If the business model cannot adapt to the new value-delivery and risk structures created by agentic work, technical maturity is unlikely to translate into scalable adoption (Juntunen, 2017). As a result, Level 2 adoption is treated in this thesis as a strategic achievement enabled by dynamic capabilities rather than as a direct consequence of technical progress.

1.8. Conceptual Model of the Thesis

This section combines the theoretical foundations and contextual conditions developed across Sections 1.1 to 1.7 into a unified conceptual model explaining the transition toward Level 2 AI adoption. The model integrates three analytical layers: technological and structural preconditions, company-level dynamic capabilities, and the platform ecosystem contexts. Together, these elements explain why advanced AI agent frameworks achieve semi-autonomous execution of business roles in some organizations while remaining experimental in others (Teece, 2007; Sohn, 2024).

At the input level, the model identifies two categories of technological and structural preconditions. First, AI agent frameworks require specific architectural capabilities to operate beyond basic automation. Prior research emphasizes that independent agent operation depends on the integration of memory mechanisms, planning functions, and tool-use capabilities, which enable agents to observe, plan, and act within complex environments (Russell & Norvig, 2021; Tang et al., 2025). These elements define the technical potential of agentic systems but do not independently determine an organization’s ability to utilize Level 2 AI agents.

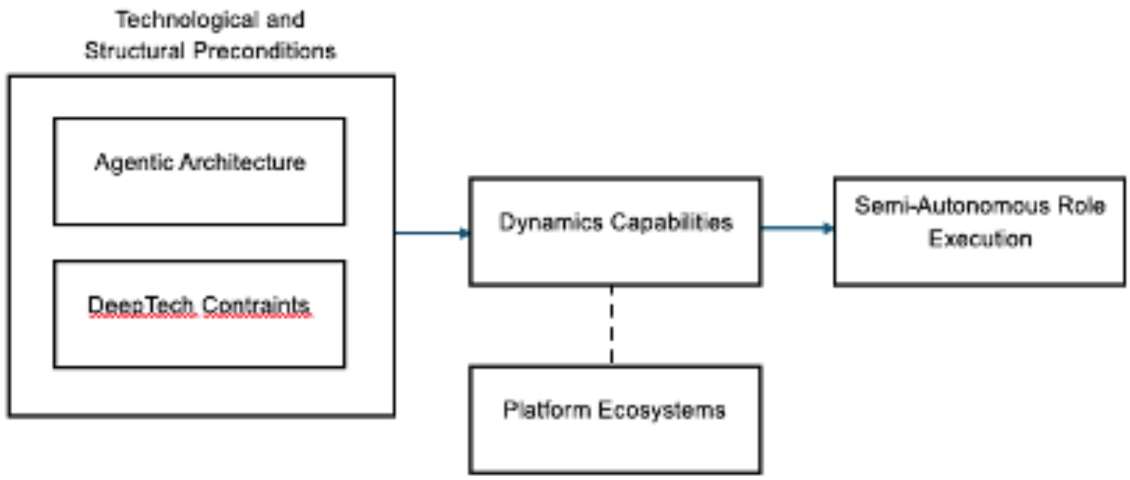
Second, the model incorporates the structural constraints that DeepTech enterprises face, which condition how this technical potential can be commercialized. These constraints include high capital intensity, dependence on advanced infrastructure, and the persistent challenges related to transforming experimental agent performance into reliable operational settings (Luitse, 2024). While such constraints do not eliminate the feasibility of Level 2 adoption, they significantly shape the companies’ speed, scope, and risk profile of scaling AI agent frameworks. This is why technical readiness must be understood together with these structural limitations rather than in isolation.

Dynamic capabilities constitute the central mediating mechanism within the model, linking technical and structural preconditions to organizational outcomes. Drawing on the sensing-seizing-transforming framework, the model helps to argue that organizations must actively interpret, mobilize, and reconfigure their resources to convert agentic potential into operational value (Teece, 2007; Warner & Wäger, 2019). Sensing enables firms to identify suitable business roles for agent deployment and to assess alignment between agent capabilities and organizational needs. Seizing involves committing resources, designing business models, and integrating AI agents into workflows. Transforming requires the reconfiguration of organizational structures, governance arrangements, and routines to support sustained human-agent collaboration. In this framework, dynamic capabilities do not represent outcomes of adoption; instead, they function as the processes that enables the adoption of Level 2 AI agents.

These enabling processes operate within platform ecosystems that shape the conditions under which dynamic capabilities can be exercised. Platform infrastructures, governance arrangements, and boundary resources influence how organizations access computational resources, deploy agents, and manage dependencies (Ghazawneh & Henfridsson, 2013; Hein et al., 2020; Engert et al., 2025). Stakeholders positioned as creators, integrators, or adopters face different constraints and opportunities depending on their ecosystem roles, which in turn affect their ability to enact sensing, seizing, and transforming activities. While platform ecosystems enable scale and access to complementary resources, they do not eliminate organizational agency; instead, they define the strategic space within which firms operate.

The outcome of this process is Level 2 AI adoption, conceptualized as the semi-autonomous execution of defined business roles by AI agents under organizational supervision, rather than full organizational automation (Sohn, 2024). As illustrated in Figure 4, the model demonstrates that scaling AI agent frameworks is neither a purely technical progression nor a linear adoption path. Instead, it emerges from the interaction between agent architecture, DeepTech constraints, dynamic managerial processes, and platform ecosystem governance.

Figure 4. *Conceptual model of the thesis*



Source: Compiled by the author

This conceptual model provides the analytical foundation for the empirical part of the thesis. It guides the investigation of how DeepTech ventures operationalize AI agent frameworks in practice and how differences in capabilities, ecosystem positioning, and governance conditions influence the transition from experimental systems to Level 2 adoption.

2. METHODOLOGICAL PART

2.1 Research Design and Philosophical Stance

This thesis uses a qualitative, exploratory research design underpinned by the interpretivist paradigm. The study investigates how AI agent technologies, particularly those demonstrating Level 2 autonomy, are adopted and scaled in businesses. Due to the emerging and context-dependent nature of this technological innovation, an interpretivist approach is best suited to examine AI agents ’adoption by enterprises. This paradigm recognizes that individuals interacting with complex systems, such as AI agents, interpret these technologies differently based on their professional roles, organizational environments, and sector-specific constraints. As emphasized by Saunders et al. (2023), interpretivism supports the exploration of subjective meanings and lived experiences, making it well-suited to examining complex socio-technical systems. The methodological positioning and reporting logic of this thesis follow the methodological guidance for Master’s theses at Vilnius University Business School (Oželienė, 2024).

Rather than attempting to identify universal patterns that can be statistically generalized to large populations, this study seeks to understand the complex realities and context-specific dynamics of AI agent adoption. The research problem involves a diverse set of stakeholders, including technology creators, integrators, and adopters, each of whom holds distinct perspectives on value delivery, risk, autonomy, and integration within their operational environments. These perspectives cannot be reduced to quantifiable variables without losing contextual meaning. Therefore, the use of an interpretivist lens is justified by the complexity of the subject matter and the diversity of stakeholder experiences that shape the phenomenon that is being investigated.

This study follows an abductive research logic, which allows for iterative movement between empirical data and theoretical frameworks. The abductive approach is particularly suitable for research on emerging technologies, where existing theories often lag behind real-world developments. As variousthemes emerge during the analysis, the researcher returns to the literature to refine, extend, or reframe theoretical understanding. This cyclical movement between theory and data supports the study’s aim of applying and, where appropriate, extending existing frameworks related to business models, governance, and dynamic capabilities within AI ecosystems.

2.2 Sampling Strategy and Participants

The sampling strategy employed in this thesis is purposive non-probability sampling, a widely accepted qualitative technique for selecting information-rich cases. The aim was not to achieve statistical representativeness but rather to gain deep, expert-level insights from individuals directly involved in different stages of AI agent implementation. As suggested by Bogner, Littig, and Menz (2009) suggest, expert interviews are particularly suitable when the research goal is to access practical insider knowledge that would not be available through other means.

Participants were selected based on their professional involvement with AI technologies, particularly in the development, deployment, and use of agent systems. The final sample had five individuals representing three distinct functional roles in the AI agent ecosystem: creators, integrators, and adopters. Creators are responsible for developing and offering AI agent infrastructure and solutions. They are experts in the systems' technical capabilities, design logic, and business model configuration. Integrators acted as the intermediaries who translated theoretical capabilities into operational deployments. They were responsible for aligning AI agent functionality with existing systems, managing workflows, and addressing integration bottlenecks. Adopters held senior decision-making roles within organizations and evaluated the business value, operational risks, and strategic fit of AI agent technologies within their respective sectors.

This role-based differentiation was crucial for triangulating perspectives and understanding the full lifecycle of AI agent implementation. It also allowed for the identification of convergent and divergent themes among participants, thereby enhancing the analytical depth of the findings. Each interview provided context-specific insights that were later analyzed for patterns, contradictions, and complementarities across the AI agent value chain. The aim was depth of understanding rather than representativeness across the ecosystem. An overview of the expert interviews, including participant roles and organizational contexts, is provided in Annex 1.

2.3 Data Collection Instrument and Process

Data were collected through semi-structured expert interviews conducted via online videoconferencing platforms in November and December 2025. The semi-structured format allowed the researcher to maintain a balance between consistency and flexibility. A core set of open-ended questions ensured that all key themes related to business models, system integration, governance, and scaling strategies were covered across interviews. Semi-structured interview guides tailored to creators, integrators, and adopters were used to ensure role-specific relevance while maintaining thematic consistency across expert interviews (Annex 2). At the same time, the format's flexibility allowed participants to elaborate on emerging issues, reveal hidden challenges, and introduce themes that were not originally anticipated. The semi-structured interview guide used to ensure thematic consistency across interviews is presented in Annex 3.

Potential logistical challenges and the geographical dispersion of participants influenced the decision to use online interviews. All participants consented to audio recording, which facilitated accurate transcription and analysis. An AI-based transcription tool, Google Gemini, was used to support the transcription of the interviews, followed by a manual review to generate the initial transcripts, which were then manually reviewed and edited by the researcher to ensure they retained the original spoken content. Particular attention was paid to technical terminology and the contextual meaning of key phrases, ensuring that the transcripts accurately represented the participants' intended messages.

The interview guide was aligned with the study's central research questions and organized into thematic blocks. These included sections on technology architecture, integration challenges, value propositions, risk evaluation, and platform dependency. While each interview followed this general structure, the researcher also enabled spontaneous conversations that resulted in insights and context-specific examples that were not initially anticipated.

2.4 Data Analysis Method

The collected data were analyzed using Reflexive Thematic Analysis (TA) as outlined by Braun and Clarke (2006). This method was chosen due to its flexibility and alignment with the interpretivist research paradigm. Reflexive TA enables researchers to identify, analyze, and report patterns (themes) within the data while acknowledging the active role of the researcher in the interpretive process. It also accommodates both deductive and inductive coding strategies, which were particularly useful for this study. The method is particularly suitable for small, expert-driven datasets where interpretation and contextual meaning are crucial.

The analysis process began with multiple readings of the transcripts to ensure data familiarization with the data. During this stage, the researcher observed the initial impressions and preliminary patterns, particularly questions regarding data quality, organizational readiness, and platform constraints. Initial codes were then generated using a hybrid approach. Deductive codes were based on theoretical constructs such as "governance mechanisms," "value capture models," and "dynamic capabilities," while inductive codes emerged from the raw data itself, including terms such as "data readiness bottleneck" and "vendor lock-in."

These codes were then arranged into broader candidate themes, which were further reviewed and refined in relation to both the coded data and the full dataset. The final themes were organized to reflect the multi-role perspectives (creator, integrator, adopter) and to answer the research questions concerning barriers, strategies, and structural conditions for AI agent scaling. The use of direct quotes and contrasting views among participants helped to illustrate these themes and provide a rich, nuanced narrative.

2.5 Ethical Considerations

This research was conducted in full accordance with Vilnius University's Code of Academic Ethics and the General Data Protection Regulation (GDPR). All participants received a clear explanation of the research aims, the use of their data, and their rights, including the right to withdraw from the research process at any point without consequence. Consent was obtained before each interview, with participants affirming their understanding and agreement to be recorded for this thesis.

Anonymity and confidentiality were maintained throughout the research process. Participants were referred to by generic role descriptors (e.g., Creator, Integrator) rather than by name or specific organizational affiliation. Sensitive business information, including financial details and client identities, was excluded from the transcripts and final write-up unless explicit permission was granted. Additionally, the thesis focused on participants' professional insights rather than personal opinions, thereby maintaining a boundary between professional roles and individual biases. Full interview transcripts were not included in the thesis due to confidentiality considerations. However, they are available upon request.

2.6 Research Quality and Trustworthiness

Given the qualitative nature of the research, the study emphasizes the criteria of trustworthiness as proposed by Lincoln and Guba and reiterated by Saunders et al. (2023). These criteria include credibility, transferability, dependability, and confirmability.

Credibility was achieved through triangulation of perspectives across ecosystem roles, drawing on insights from multiple stakeholder roles to ensure that findings were not overly dependent on any single perspective. Participants confirmed the accuracy of summarized insights during or after the interviews, a process like member checking. Transferability was enhanced through the provision of detailed descriptions, including comprehensive accounts of participant contexts, organizational environments, and sector-specific challenges. This level of detail allows readers to assess whether and to what extent the findings may apply to other settings.

Dependability was ensured by maintaining detailed records of the research process, including the development of the interview guide, the transcription and coding procedures, and the analytical decisions made at each stage. Confirmability was supported by direct quotations and the consistent application of the thematic analysis framework, which allowed the findings to emerge from the data rather than from the researcher's personal assumptions.

Reflexivity was an essential component of the research process. The researcher acknowledged their own background in digital technologies, which could influence the interpretation of participants' statements. To mitigate this risk, the analysis emphasized contrasting perspectives and highlighted participants' own words and thoughts wherever possible.

2.7 Methodological Limitations

This study has several limitations that must be acknowledged. First, the use of purposive sampling and a small number of participants limits the generalizability of findings to broader populations. The insights derived are context-specific and reflect the experience of individuals working within industries and organizational structures. While analytical generalization is possible, caution must be exercised in applying these findings to other sectors or technological domains.

Second, the reliance on expert interviews means that the research captures only those aspects of AI agent implementation that are evident or salient to professionals. Other organizational factors, such as internal politics, informal routines, or unarticulated cultural norms, may also influence AI adoption processes but were beyond the scope of this study. In addition, the rapid pace of progress in AI technology means that some of the technical challenges and strategic responses identified in this research may quickly become outdated.

Despite these limitations, the study provides valuable exploratory insights into the mechanisms of AI agent adoption and scaling from a multi-stakeholder perspective. It contributes to the broader academic and practical understanding of how DeepTech ventures navigate the challenges of business model innovation and ecosystem governance in the context of emerging intelligent technologies. Even though the study employed a qualitative research design, triangulation was achieved by including multiple ecosystem roles (creators, integrators, and adopters), enabling comparison of perspectives across positions in the AI agent value chain.

3. Findings

This chapter presents the empirical findings based on the from five semi-structured expert interviews conducted with professionals directly involved in the development, integration, and adoption of AI agent technologies. The participants represent three distinct ecosystem roles: one Creator (a solutions architect responsible for agent infrastructure and system design), one Integrator (an engineering lead accountable for deployment and system integration), and three Adopters (senior product and technology decision-makers from FinTech, e-commerce, and marketing startups).

The purpose of this chapter is to report on how AI agent frameworks are currently implemented, evaluated, and scaled in organizational settings, with particular attention being paid to systems approaching Level 2 adoption, in which agents perform defined business tasks under human supervision. The findings reflect participants' practical experiences and perceptions rather than theoretical expectations.

The analysis identifies five cross-cutting empirical themes: (1) data readiness and quality, (2) human oversight requirements, (3) integration and interoperability constraints, (4) governance, security, and compliance considerations, and (5) strategic value and return-on-investment (ROI) assessment. Each theme is reported by differentiating perspectives across ecosystem roles to capture similarities and differences in how AI agent adoption is experienced in practice.

3.1 Data Readiness and Quality

Across all interviews, data readiness emerged as the most critical prerequisite for the successful deployment of AI agents. Participants consistently emphasized that agentic systems are fundamentally constrained by the structure, quality, and contextual richness of the data they have access to. Regardless of the sophistication of the underlying models, inadequate data preparation was described as something that limits both system reliability and organizational trust.

Interviewees repeatedly referred to the principle of “garbage in, garbage out” when discussing agent performance. The principle outlines that poor-quality data produces flawed results, and in the case of AI agents, can affect the reliability of agents. Integrators and adopters noted that data issues are rarely visible at the beginning of projects and often surface only once agents begin operating within real workflows. One integrator highlighted that insufficient attention to data quality during early discovery phases frequently leads to a project’s failure, as agents amplify existing data inconsistencies rather than compensating for them. From the adopter's perspective, limitations in historical data-collection practices were described as a structural obstacle that agents cannot resolve on their own.

Beyond basic cleanliness, participants stressed the importance of contextual completeness. Several adopters explained that agents lack the implicit business understanding that human workers rely on when interpreting data. As a result, even technically correct datasets may produce misleading outcomes if critical contextual markers are absent. For example, adopters described scenarios where agents were unable to distinguish between different operational categories - such as advertising formats or product attributes - because the underlying datasets did not encode these distinctions explicitly. In such cases, agents produced outputs that were formally valid but operationally incorrect.

To mitigate these issues, organizations reported investing substantial manual effort into data preparation. Rather than relying on centralized or universal data repositories, adopters described preparing datasets on a per-use case basis. This process included extracting data from multiple sources, validating consistency, and adding explanatory metadata to ensure that agents could interpret the information correctly. One adopter characterized this as recurring overhead, noting that each new agentic use case required a bespoke data-preparation process rather than the reuse of existing pipelines.

Participants also highlighted temporal instability as a data-related challenge. Even when data was initially prepared to a sufficient standard, ongoing changes in upstream systems, business logic, or data sources could quickly degrade agent performance. This required continuous monitoring and maintenance, further increasing the operational burden associated with agent deployment. As a result, data readiness was not perceived as a one-time prerequisite but as an ongoing operational responsibility.

Overall, the findings indicate that data readiness functions as a foundational constraint on agent adoption. Organizations do not perceive agent failures primarily as model errors but rather as symptoms of deeper limitations in the data infrastructure. Consequently, decisions about whether and where to deploy AI agents are heavily influenced by assessments of data availability, structure, and maintainability, often preceding considerations of model capability or automation potential.

3.2 Human Oversight in Agentic Workflows

Despite widespread interest in autonomous AI agents, all participants emphasized that human involvement remains essential to current agentic workflows. Interviewees did not see human oversight as a temporary safeguard during early experimentation. They consistently described it as a necessary operational requirement for maintaining system reliability, organizational trust, and accountability at least over the short term.

The primary driver of continued human involvement is the perceived risk associated with unsupervised agent behavior. Several adopters cited prior negative experiences in which AI-driven automations produced incorrect or harmful outcomes when left unsupervised. These incidents included data corruption, misclassification, and unintended negative financial consequences. As a result, organizations implemented explicit approval mechanisms that required human validation before agent outputs were executed or remained in production systems. One adopter described these controls as “man-in-the-middle” safeguards designed to prevent irreversible errors.

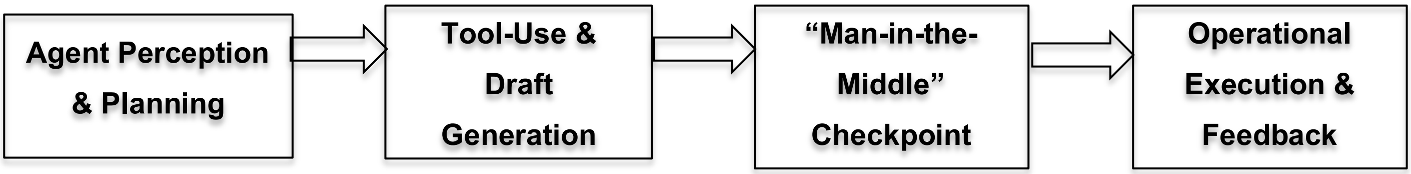
Trust was repeatedly identified as a limiting factor. Participants reported that, while agents can generate outputs that appear correct, the underlying reasoning process is often unclear, making it difficult for organizations to delegate full responsibility to AI agents with confidence. Creators and integrators noted that agents frequently encounter edge cases that were not anticipated during development or testing, reinforcing the need for human judgment in non-standard situations. Human oversight was described as a verification mechanism and a diagnostic tool for identifying failure modes and improving system design.

Regulatory and compliance considerations further reinforced the necessity of human involvement. Adopters operating in tightly regulated industries reported that fully automated decision-making is either restricted or explicitly prohibited by law. In these contexts, human validation is mandatory regardless of technical capability. Participants emphasized that even highly accurate agents cannot bypass these requirements, as accountability must ultimately remain with a human who makes the final decisions.

Participants also described autonomy as a gradual and conditional process rather than a binary transition. Organizations reported increasing agent responsibility incrementally as confidence in system performance grew. However, this progression depended on evaluation mechanisms that enabled organizations to monitor outcomes and detect deviations. Without such mechanisms, participants indicated that granting greater autonomy would expose organizations to unacceptable operational risk.

Overall, the findings suggest that current AI agent deployments employ a hybrid model that combines automated execution with human supervision. Human involvement is not framed as a failure of agent capability but as an operational necessity shaped by trust, risk management, and regulatory considerations. This persistent reliance on human oversight significantly influences how organizations scope agent responsibilities and defines the practical limits of autonomy in real-world deployments. This operational configuration, where agentic execution is moderated by mandatory human validation to ensure reliability and accountability, is illustrated in the Supervised Autonomy Workflow in Figure 5.

Figure 5. *The supervised autonomy workflow: human-in-the-loop integration*


Agent Perception & Planning
Tool-Use & Draft Generation
"Man-in-the-Middle" Checkpoint
Operational Execution & Feedback

Source: Compiled by the author

3.3. Integration and Interoperability Constraints

The participants stated that integration and interoperability are also among the most persistent operational challenges in scaling AI agent frameworks. Interviewees consistently described the ability of agents to interact with existing technical systems as a key condition for moving beyond isolated pilots into sustained production use.

Across interviews, dependency on APIs emerged as a baseline requirement for agentic workflows. Adopters emphasized that automation becomes impractical when key systems do not result in stable, well-documented interfaces. APIs were framed as an optimization feature but also as the minimum technical precondition for enabling agents to access data, trigger actions, and complete workflows within organizational systems.

Legacy infrastructure was also identified as a recurring constraint. Participants described situations in which internal AI capabilities were not the bottleneck, yet automation stalled because external partners relied on outdated enterprise systems. These partner-side systems lacked the protocols necessary to support agent-driven interaction, thereby preventing agents from executing routine operational tasks, such as inventory checks, order placement, or information synchronization across organizations. As a result, interoperability limitations were experienced not only as an internal IT issue but as an ecosystem-level problem shaped by the technological maturity of suppliers and partners.

Participants further noted that the broader tooling landscape remains unstable. Adopters reported that many third-party tools and services in the AI agentic ecosystem are at the prototype or early minimum viable product stage, making them difficult to deploy safely and at scale in production environments. These tools were characterized as unreliable, inconsistently documented, or lacking security assurances, which only increased the integration workload and limited the set of solutions that organizations would consider viable for operational deployment. Creators similarly acknowledged that rapid innovation often outpaces reliability, increasing the burden on organizations to test, monitor, and limit agentic systems during integration.

A related theme, concerning infrastructure control and platform dependence, but also reappeared during interviews. Creator described implementing an agentic workflow using an open-source agent development kit alongside Gemini and Google Cloud Vertex AI, illustrating that production deployment is often anchored in hyperscaler ecosystems and managed external services. At the same time, Adopters expressed a preference for reducing long-term dependence on “black box” providers by retaining implementation control through self-service APIs and open or semi-open frameworks. This preference was described as a way to preserve flexibility, manage vendor risk, and maintain control over operational infrastructure as agentic workflows expand.

Overall, the findings show that integration constraints extend beyond technical compatibility. Interoperability depends on internal system maturity, the readiness of external partners, and the stability of the surrounding tool ecosystem. These constraints shape how organizations scope agentic initiatives in practice, often restricting deployments to narrowly defined workflows where dependencies can be controlled and risk can be managed.

3.4 Governance, Security, and Compliance

Governance, security, and regulatory compliance also emerged as some of the gatekeeping factors that powerfully shape the adoption and scaling of AI agent frameworks. During the interviews, participants emphasized that even technically viable agentic solutions can be delayed or blocked entirely if they fail to meet organizational or regulatory governance requirements.

Adopters operating in regulated industries described security approval processes as a critical bottleneck. In these contexts, internal security teams, led by Chief Information Security Officers (CISOs) - were reported to hold ultimate decision-making authority over whether AI systems could be deployed. Participants explained that this authority is often exercised with caution, with security concerns overriding potential efficiency gains. Thus, adoption timelines were described as long and uncertain, even when pilot results were positive.

Data protection and access control were outlined as concerns when it came to governance discussions. Adopters emphasized the necessity of strict data segmentation to prevent unintended data exposure across clients, systems, or workflows. One participant described categorizing data into different sensitivity tiers, with specific categories, like financial records, being explicitly excluded from interaction with AI systems. Other data types, such as descriptive or publicly available information, were viewed as lower risk and more suitable for agentic processing. These practices shaped which use cases could be automated and constrained the scope of agent deployment.

Participants also highlighted security risks associated with third-party tools and emerging vendors. Several adopters expressed skepticism about newly released AI services, citing insufficient documentation, unclear data-handling practices, and unresolved security vulnerabilities. This uncertainty reinforced a preference for cautious adoption strategies, such as limiting agent access to non-critical systems or requiring extensive internal review before integration into operations.

Governance requirements also influenced architectural choices. Organizations often favored solutions that allowed high-precision control over data flows, auditability, and intervention mechanisms. Human approval checkpoints were frequently embedded into workflows as a governance safeguard, ensuring accountability even when agents performed substantial portions of the task. These controls were not seen as optional enhancements but as mandatory conditions for operating agents within production environments.

Overall, the findings demonstrate that governance and security concerns also function as structural constraints rather than minor potential problems. Adoption decisions are shaped less by what agents are technically capable of doing and more by whether organizations can ensure compliance, accountability, and risk containment. These constraints significantly narrow the range of acceptable use cases and contribute to the cautious, incremental nature of agentic adoption observed across participants' ventures.

3.5 Strategic Value and ROI Assessment

Across all interviews, participants emphasized that the ROI of AI agent deployments remain highly uncertain and context-dependent. Unlike traditional software projects, where costs and benefits can be estimated with relative precision, agentic systems introduce probabilistic behavior and evolving performance characteristics that complicate the development of a formal business case. As a result, organizations rely on a combination of proxy metrics, incremental experimentation, and strategic judgment rather than formal financial models.

Several adopters highlighted that ROI is rarely calculated through direct cost-benefit analysis. Instead, decision-making is guided by ancillary indicators such as operational speed, fault tolerance, and workload substitution. One adopter described setting concrete performance thresholds, noting that if an agent fails to exceed a predefined accuracy level, the initiative is terminated regardless of its theoretical potential. Another adopter framed value primarily in terms of labor efficiency, assessing how many employees would otherwise be required to perform the same review or categorization tasks. These approaches indicate that ROI is retrospectively evaluated, based on the observed system behavior, rather than prospectively through predictive modeling.

Uncertainty regarding agent reliability further limits investment decisions. Participants expressed concern that agents may perform well in isolated scenarios but do worse under scale or when exposed to edge cases. This uncertainty makes organizations hesitant to commit to large-scale deployments of AI agents. Instead, they favor tightly scoped pilots that limit operational exposure and financial risk. One integrator described this approach as deliberately “slicing” workflows into smaller components, allowing organizations to validate agent performance incrementally before extending their scope or responsibility.

In the absence of clear financial justification, strategic considerations play a decisive role in adoption decisions. Adopters acknowledged that some investments in agentic systems are driven by expectations about future competitiveness rather than immediate returns. These decisions were described as “strategic bets” or “gut feelings,” reflecting the belief that early experimentation is necessary to avoid technological lockout or capability gaps in the future. However, such strategic motivation does not eliminate caution with the motivation instead reinforcing a preference for easily reversible and modular implementations.

Taken together, the findings suggest that ROI assessment for AI agents is an ongoing evaluative process rather than a one-time calculation. Organizations continuously reassess value as agents are tested, adjusted, and used within operational environments. This incremental and cautious approach reflects both the novelty of agentic systems and the high perceived cost of failure, shaping how and where organizations are willing to experiment with autonomous execution.

4. Analysis and Interpretation of the Research Results

The primary aim of this thesis is to analyze the entrepreneurial challenges and opportunities associated with scaling AI agent frameworks toward Level 2 business adoption, referred to as AI equalization. While academic literature increasingly conceptualizes LLM-based agents as a paradigm shift from static prediction toward autonomous task execution (Russell & Norvig, 2021; Tang et al., 2025), the empirical findings of this study suggest that the transition from experimental agentic systems to reliable organizational roles remains constrained, uneven, and highly depended on context in enterprises.

This chapter moves beyond the descriptive presentation of findings and offers an interpretative analysis based on in three complementary theoretical perspectives. First, the Dynamic Capabilities framework (Teece, 2007; Warner & Wäger, 2019) is used to examine how organizations sense, seize, and attempt to use emerging AI agentic technologies to create operational value. Second, the platform ecosystem governance theory (Hein et al., 2020; Luitse, 2024) provides context to understanding how infrastructural dependencies, boundary resources, and power asymmetries influence entrepreneurial decision-making. Third, AI adoption taxonomies (Sohn, 2024) are utilized to interpret the gap between technical capability maturity and the actual delegation of organizational roles to AI systems.

In addition to academic theory, the analysis places the empirical results next to selected contemporary industry benchmarks and users' reports, including large-scale enterprise studies and ecosystem analyses (Google Cloud, 2025; McKinsey & Company, 2024; ServiceNow, 2025; RAND, 2024). These sources are not treated as empirical evidence or academic literature but rather as contextual reference points that help assess whether the observed organizational constraints align with or diverge from dominant industry narratives surrounding agentic AI adoption.

By integrating theoretical frameworks with empirical insights and industry context, this chapter explains why the adoption of agentic AI frequently ends at the pilot implementation phase. The analysis highlights how organizational readiness, data governance structures, and platform dependency jointly shape the feasibility of AI Equalization, revealing that the primary barriers to Level 2 adoption are both technological limitations and structural and managerial conditions present within enterprises.

4.1. The Gap Between Technical Maturity and Business Adoption

The theoretical framework developed in Chapter 2 described a progressive relationship between technical maturity and organizational adoption. Specifically, contemporary AI agents equipped with persistent memory, planning capabilities, and tool use are theoretically capable of reaching Level 3, which in turn enables Level 2 organizational adoption or as AI equalization - where agents perform defined business roles independently while still under supervision (Tang et al., 2025; Sohn, 2024). However, the empirical findings of this study challenge the assumption that technical readiness translates directly into organizational deployment.

Across all interviews, participants consistently pointed out that the transition from technically capable agents to operationally trusted systems is neither linear nor automatic. While interviewees acknowledged that current AI agent architectures increasingly demonstrate advanced planning and tool management capabilities, these technical attributes alone were insufficient to justify role-level delegation within organizations. Instead, adoption stalled when it came to measuring technical performance and organizational risk tolerance. This suggests that Level 3 technical maturity is a necessary but not a sufficient condition for businesses to adopt Level 2 AI agents.

A central empirical insight concerns the distinction between functional autonomy and organizational legitimacy. Agents may be capable of executing tasks independently in tightly controlled environments, yet organizations remain reluctant to assign them responsibility for business-critical processes. Doubts about model intelligence do not primarily drive this reluctance since the concerns relate to data reliability, error propagation, accountability, and governance. In practice, organizations treat autonomy with with flexibility rather than a technical threshold, incrementally extending agent responsibility only after repeated validation cycles.

The findings further reveal that organizational adoption is shaped by the weakest element in the socio-technical system, not by the pinnacle of technical capability. Even when agents demonstrate robust performance when doing isolated tasks, flaws in data structures, integration interfaces, or oversight mechanisms prevent the agents' escalation to role-level autonomy. This aligns with the “wooden barrel” theory pointed out above, whereby overall system performance is constrained by its least developed component (Tang et al., 2025). As a result, organizations keep using hybrid configurations that resemble advanced augmentation rather than genuine role equalization.

From an entrepreneurial perspective, this gap reframes the scaling challenge faced by DeepTech ventures. Rather than focusing exclusively on improving agent intelligence, ventures must also address organizational enablers that determine whether autonomous agents can be trusted, governed, and economically justified. The empirical evidence indicates that the decisive barrier to Level 2 adoption lies not in agent ability, but in the institutional conditions under which agents are used.

Taken together, these findings suggest that the relationship between technical maturity and business adoption is best understood as decoupled but interdependent. Technical advances expand the space of possible adoption, but organizational constraints determine the actual boundaries of deployment. This gap forms the analytical foundation for the subsequent sections, which examine how data readiness, human oversight, platform governance, and dynamic capabilities mediate the translation of agentic potential into operational reality.

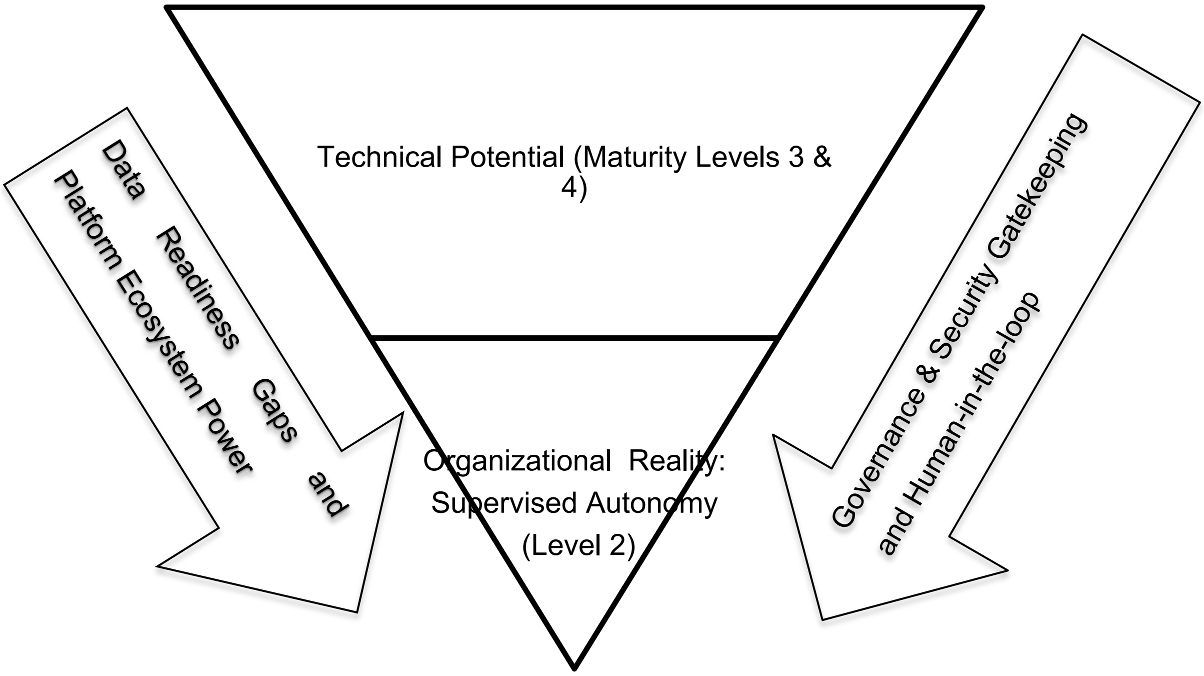
Furthermore, the empirical findings suggest that the gap between technical maturity and organizational adoption is primarily shaped by organizational preconditions rather than by deficiencies in agent intelligence. Data readiness and the persistence of human oversight emerge as mutually reinforcing constraints that prevent the transition from technically autonomous systems to role-level delegation in enterprise settings.

Participants consistently emphasized that agent autonomy is constrained by the quality, structure, and contextual completeness of enterprise data. While contemporary agents may technically possess advanced planning and tool-use capabilities, their operational effectiveness is undermined by fragmented data infrastructures that were not designed for autonomous execution. This finding reframes the theoretical “sim-to-real” gap identified by Tang et al. (2025) as a data-to-context gap, in which agents fail not because cognitive limitations but because they lack access to reliable, semantically structured inputs required for consistent decision-making. Extending the DeepTech perspective outlined in Section 2.4, the findings suggest that data readiness accounts for a significant amount of capital intensity in AI agent ventures. Significant engineering effort is also required to cleanse, contextualize, and expose organizational data in a manner suitable for agentic operation, creating a substantial barrier to commercialization.

The persistence of human-in-the-loop (HITL) arrangements further reflects organizational responses to this data uncertainty. Although Sohn (2024) conceptualizes Level 2 adoption as role-level autonomy, empirical evidence from this study indicates that enterprise implementations remain hybrid. Even when agents are capable of executing tasks independently, organizations retain human oversight to manage liability, ensure compliance, and mitigate the risks associated with opaque model behavior. Rather than serving as a transitional phase toward full automation, HITL emerges as a structural governance mechanism that compensates for unresolved data quality issues and trust deficits.

This finding challenges a strict interpretation of AI equalization as the fully autonomous execution of roles. Instead, Level 2 adoption in practice resembles a collaborative configuration in which agents perform operational work while humans retain accountability and decision authority. This aligns with prior conceptualizations of AI as a collaborator rather than a replacement actor and reflects broader concerns related to black-box decision-making and organizational risk exposure (Black et al., 2024; Kilian, 2025). Consequently, the analysis suggests that HITL should be understood not as a temporary limitation, but as a stable feature of Level 2 adoption in high-stakes organizational contexts. This consolidation of high technical potential and persistent organizational constraints, which results in the narrowed reality of supervised autonomy, is visualized in the Maturity-Adoption Bottleneck model as per Figure 6.

Figure 6. *The maturity-adoption bottleneck in AI agent scaling*


Technical Potential (Maturity Levels 3 & 4)
Organizational Reality: Supervised Autonomy (Level 2)
Data Readiness Gaps and Platform Ecosystem Power
Governance & Security Gatekeeping and Human-in-the-loop

Source: compiled by the author

Taken together, data readiness and human oversight form a coupled constraint that explains why technical maturity does not automatically translate into business adoption of AI agents. Organizations extend agent autonomy only to the extent that data infrastructures and governance mechanisms allow risks to be managed. This insight reinforces the argument that scaling AI agent frameworks is fundamentally an organizational challenge, in which technological capability expands possibilities, but internal business conditions determine realizable outcomes within DeepTech ventures.

4.2. Dynamic Capabilities in the Agentic Era

The Dynamic Capabilities framework (Teece, 2007; Warner & Wäger, 2019) was used to interpret how organizations attempt to scale AI agent frameworks beyond experimental pilots toward Level 2 adoption. The empirical findings confirm the continued relevance of sensing, seizing, and transforming capabilities, but also reveal that these capabilities appear differently in the context of agentic AI compared to traditional digital transformation initiatives.

Rather than functioning as sequential or clearly separable processes, the capabilities appear tightly interdependent and constrained by organizational and ecosystem limitations.

In the context of AI agents, sensing extends beyond the identification of technological opportunities and encompasses the systematic assessment of risk and strategic prioritization. Interviewees spoke about the fact that organizations actively avoid deploying agents in business-critical processes unless value creation can be demonstrated under unpredictable conditions. This aligns with the notion of “ancillary opportunity spotting” proposed by Hafeez et al. (2025b), where viable use cases are selected not for their technical novelty but for their alignment with broader organizational goals and risk tolerance. The empirical evidence suggests that sensing failures often lead to prolonged experimentation at scale, reinforcing the phenomenon described in the literature as “pilot purgatory” (Sjödin et al., 2023). Consequently, sensing in the agentic era is less about discovering where AI can be applied and more about determining where it should be applied.

The seizing of agentic opportunities further illustrates the complexity of transforming insight into operational value. While Warner and Wäger (2019) conceptualize seizing as rapid resource mobilization and prototyping, the findings indicate that deploying AI agents requires extensive co-creation and integration across organizational and ecosystem boundaries. Participants described significant challenges related to interoperability, legacy-system rigidity, and API dependence, confirming that deriving value from AI agents depends on the ability to integrate diverse technical infrastructures. This supports Hafeez et al.’s (2025b) argument that business model co-creation capabilities are essential in complex digital ecosystems, as value realization depends on aligning technical feasibility with organizational workflows and partner systems. In practice, seizing is frequently stalled not by lack of intent, but by structural constraints embedded in existing IT architectures (Tupe & Thube, 2025).

The transforming capability emerges most clearly when examining governance and organizational reconfiguration. Consistent with Teece’s (2007) framework, scaling AI agents requires changes to internal structures, decision-making delegations, and accountability mechanisms. Empirical findings highlight that organizations attempting to move beyond Level 1 augmentation must develop new management arrangements, including formal oversight mechanisms, data access controls, and escalation procedures for agent failures. These arrangements function as a critical dynamic capability, enabling firms to manage the probabilistic and unclear nature of LLM-based systems. Without the ability to transform governance structures accordingly, organizations could be unable to delegate responsibility to agents, reinforcing the need of human oversight and limiting adoption to assistive use cases. This observation supports the view that governance itself constitutes a dynamic capability necessary for mitigating the structural risks associated with AI deployment (Kilian, 2025).

Taken together, the findings suggest that dynamic capabilities in the agentic era are not about speed of transformation but more about coordination across technical, organizational, and governance domains. The inability to align sensing, seizing, and transforming activities under conditions of uncertainty explains why many organizations fail to scale agentic systems despite their technical readiness. Dynamic capabilities, therefore, function as the primary mechanism through which firms attempt - often imperfectly - to bridge the gap between agentic potential and organizational adoption.

The theoretical framework also emphasizes the relevance of platform-dependent entrepreneurship and infrastructural power in AI-driven ecosystems, where ventures and adopters rely on external platform owners for access to essential resources (Yu & Sekiguchi, 2024; Luitse, 2024). The empirical evidence of this study is consistent with these perspectives, suggesting that platform dependence is not merely a background condition but a structural factor shaping how AI agent frameworks can be scaled toward Level 2 adoption.

This dependence is particularly critical for resource-constrained companies or startups, where platforms function as an enabling route to access AI capabilities that would otherwise require substantial internal investment. Research on B2B SMEs shows that AI platforms often become a practical pathway for integrating AI technologies into operations (Wei & Pardo, 2022). In this study, a similar rationale appears in interviews, which describe agent deployments being hosted in external cloud tools and managed services, with the rationale also showing up in adopters' emphasis on retaining implementation control as reliance on these tools and services deepens.

Across interviews, participants described that production deployment of agentic workflows is commonly depends on cloud ecosystems that provide access to foundation models, compute, and development services. In the empirical part of the thesis, this was illustrated through references to managed AI services and tooling ecosystems used in real deployments. This pattern aligns with Luitse's (2024) argument that platform owners exercise infrastructural power through vertical integration and abstraction, packaging complex operations into standardized services that accelerate development while constraining adopters and ventures within proprietary stacks. From this perspective, platform dependence functions as both an enabler of rapid experimentation and a limit on strategic freedom, because access to models, APIs, and compute is described in platform-defined technical and contractual terms (Luitse, 2024; Canboy & Khlif, 2025).

The findings also imply that platform power is expressed through governance mechanisms within the architecture. Access to critical capabilities is operationalized through boundary resources, such as APIs and SDKs, which structure the relationship between platform owners and complementors by defining the types of interactions possible and the constraints under which they occur (Ghazawneh & Henfridsson, 2013). In practice, participants framed APIs as a minimum requirement for integration and automation, but they also emphasized that security requirements and internal approval processes often constrain adoption. This supports the interpretation that platform participation is governed not only by technical feasibility but also by compliance rules and permission structures that can override implementation readiness.

A key empirical nuance is that platform dependence is not limited to external cloud providers; it also encompasses control over implementation choices. Adopters expressed a preference for self-service APIs and open or semi-open frameworks where possible, explicitly framing this as a strategy to reduce reliance on “black box” solutions and to preserve flexibility as agentic workflows expand. This orientation reflects a practical concern with the risk of lock-in. When agents become embedded in operational processes, the costs of switching to another supplier rise, and the organization’s autonomy becomes increasingly reliant on a provider’s decisions. In this sense, the empirical material complements the concept of platform-dependent entrepreneurship by indicating that dependency is experienced not only through reliance on distribution channels but also through reliance on technical infrastructure and governance arrangements that are outside the business’s direct control (Yu & Sekiguchi, 2024).

Taken together, the findings suggest that scaling AI agent frameworks toward Level 2 adoption involves a strategic navigation of platform dependence rather than a purely technological scaling problem. Platform ecosystems provide essential resources that enable deployment and integration, yet they also limit ventures through governance mechanisms, security constraints, and infrastructural dependencies. For DeepTech entrepreneurs, this implies that progress toward AI equalization depends not only on improving technical capabilities and organizational readiness, but also on deliberate choices about architecture, interoperability, and dependency management within platform-dominated environments.

4.3. Implications for the Conceptual Model

The conceptual model developed in Section 2.7 proposed that the transition toward Level 2 AI adoption is shaped by the interaction between technological enablers, DeepTech constraints, dynamic capabilities, and platform context. The empirical findings largely support the model, while also indicating the need for several refinements to better reflect organizational realities observed in practice.

First, the findings suggest that data readiness should be elevated within the model from a background DeepTech constraint to a primary one. While the original model treated technological enablers such as memory, planning, and tool use as the core drivers of AI agentic capability, the empirical evidence demonstrates that these capabilities remain largely ineffective in the absence of curated, structured, and context-rich organizational data. In practice, data readiness determines whether technical agent architectures can function reliably, positioning it as a foundational precondition rather than a secondary constraint.

Second, the mediating role of dynamic capabilities can be confirmed but requires further refinement. In particular, the transforming capability must explicitly encompass what can be described as governance engineering. The findings show that scaling AI agents beyond experimental pilots requires organizations to design and institutionalize new management structures, including approval workflows, access controls, and accountability mechanisms. These activities go beyond basic organizational reconfiguration and represent a distinct capability related to managing probabilistic systems and AI-related risk. Without the governance-oriented transformation, firms struggle to operationalize agentic systems at scale.

Finally, the empirical results require a refinement of the outcome dimension of the model. While the theoretical framework conceptualized Level 2 adoption as AI equalization, the findings indicate that full role autonomy is rarely realized in practice. Instead, organizations consistently deploy AI agents in supervised operational environments, where agents execute substantial portions of business tasks but remain subject to human oversight. As such, Level 2 adoption is more accurately characterized as supervised autonomy rather than complete role substitution. This adjustment reflects the persistent influence of trust, liability, and regulatory considerations on organizational adoption decisions.

Together, these refinements do not undermine the original conceptual model but rather enhance its explanatory precision. The revised model emphasizes that successful scaling of AI agent frameworks depends not only on technical maturity but also on data readiness, governance-focused transformation capabilities, and an empirically underpinned understanding of autonomy as a managed organizational condition. This adjusted framework provides a more realistic basis for interpreting how DeepTech ventures navigate the transition from technical feasibility to sustainable business adoption of AI agents

4.4. Comparative Analysis with Industry Benchmarks

The empirical findings of this study can be further contextualized by comparison with recent industry benchmark reports on enterprise AI adoption. While these industry sources do not constitute empirical evidence for this research, they provide a helpful reference point for assessing whether the challenges observed in the interviews reflect isolated organizational issues or broader systemic patterns.

The prominence of data readiness as a limiting factor in this study is consistent with industry-wide assessments of enterprise AI maturity. Recent benchmark reports indicate that, despite rapid advances in agentic and generative AI technologies, many organizations struggle to commercially operate these systems due to fragmented data infrastructures and insufficient governance capacity (Kotu et al., 2025). This reflects the empirical observation that technical agent capabilities often outpace organizational preparedness, particularly with respect to data structures, contextualization, and maintenance. Contraints related to data readiness identified among DeepTech ventures trend a broader industry condition rather than a niche technical problem.

Furthermore, the persistence of human oversight identified in this study aligns with emerging industry perspectives on the responsible deployment of AI. While industry reports frequently highlight increasing experimentation with AI agents, they also emphasize the necessity of human supervision, guardrails, and staged deployment models (Google Cloud, 2025). Analytical reports further frame AI systems as socio-technical constructs in which human judgment remains essential for managing uncertainty, accountability, and risk (RAND, 2024). This industry context supports the empirical finding that organizations rarely pursue full autonomy in practice and instead adopt hybrid operating models in which agents execute tasks under continuous human supervision.

Third, the findings related to platform dependency and infrastructural control can be placed within broader industry discussions on cloud service providers’ consolidation and vendor reliance. Industry benchmarks indicate that organizations achieving higher returns from AI investments often centralize data and AI capabilities within dominant platform ecosystems to accelerate integration and governance (Kotu et al., 2025). At the same time, policy-oriented analyses caution that such consolidation increases exposure to vendor lock-in, governance asymmetries, and strategic dependency on private infrastructure providers (RAND, 2024). The empirical findings of this study reflect this tension at a company level, where reliance on platform ecosystems enables technical scaling while simultaneously constraining strategic autonomy.

The comparison with industry benchmarks suggests that the main challenges identified in this thesis, namely data readiness, supervised autonomy, and platform dependency, are not transitional anomalies but structural characteristics of the current agentic AI landscape. These parallels reinforce the relevance of the conceptual model developed in this research and indicate that the constraints faced by DeepTech ventures are present within broader industry dynamics that shape how AI agents are adopted and managed at scale.

CONCLUSIONS AND RECOMMENDATIONS

Conclusions

This thesis examines the entrepreneurial challenges associated with scaling AI agent frameworks toward Level 2 business adoption, referred to as AI equalization, within DeepTech enterprises. By combining a multi-layered theoretical framework with qualitative empirical evidence from expert interviews, the study aimed to move from abstract discussions of agent autonomy, and to understand how such systems are implemented, governed, and scaled in organizational settings.

The findings demonstrate that the central obstacles to scaling AI agent frameworks are not primarily technological. While contemporary LLM-based agents exhibit advanced capabilities in planning, memory, and tool use, these capabilities do not directly result in organizational adoption. Instead, the transition from experimental pilots to operational business roles is constrained by structural and organizational conditions that shape how agents can be deployed in practice.

Among these conditions, data readiness emerged as the most critical prerequisite for agentic scaling. The study shows that fragmented, poorly structured, or context-deficient data environments severely limit agent reliability, regardless of model sophistication. As a result, data preparation and maintenance become continuous, resource-intensive business activities rather than preliminary technical steps. These finding challenges technology-centric narratives of AI adoption by demonstrating that data infrastructure, rather than algorithmic capability, frequently determines whether agents can even operate at all within organizational environments.

The research further reveals that Level 2 AI adoption rarely results in full role substitution in practice. Instead of delegating complete autonomy to agents, organizations consistently adopt arrangements characterized by supervised autonomy. In these configurations, AI agents can execute defined tasks and workflows, but humans keep oversight, accountability, and final decision-making authority. This persistent human involvement is driven by the lack of trust, regulatory requirements, and liability concerns, rather than by an absence of technical capability. AI equalization should be deemed as an organizational compromise between efficiency gains and risk management, rather than as a purely technical milestone.

In addition, the study highlights the central role of platform ecosystems in shaping entrepreneurial outcomes. DeepTech ventures and adopting organizations depend on external cloud infrastructures, foundation models, and platform governance mechanisms that both enable experimentation and impose constraints. While platform ecosystems lower entry barriers and accelerate technical development, they also create asymmetrical power relations that limit strategic autonomy and expose ventures to risks related to infrastructural control. These dependencies influence not only technical architecture choices but also the long-term viability of the business model.

The conclusions of this thesis suggest that scaling AI agent frameworks is fundamentally a socio-technical process. Technical maturity is a necessary but insufficient condition for Level 2 adoption. Sustainable commercialization requires alignment between agent capabilities, data infrastructures, governance mechanisms, and ecosystem conditions. Understanding this alignment is essential for explaining why many agentic initiatives remain confined to pilots despite rapid advances in AI technology.

Recommendations

Based on the findings of this study, several practical recommendations can be made for entrepreneurs, adopting organizations, and ecosystem stakeholders involved in the development and deployment of AI agent frameworks.

For DeepTech entrepreneurs, the results imply that early-stage investment priorities should extend beyond model performance and agent architecture. Particular emphasis should be placed on assessing data readiness within target organizations and designing solutions that explicitly account for potential data fragmentation and contextual gaps. Entrepreneurs should anticipate that data preparation and integration will represent a substantial portion of deployment effort and should incorporate these realities into product design, pricing models, and implementation strategies.

Organizations seeking to adopt AI agents should approach autonomy as a managed organizational condition rather than an end state. Instead of pursuing full automation, firms are advised to design workflows that keep human oversight, clear accountability structures, and staged escalation mechanisms. Incremental expansion of agent responsibilities, combined with continuous evaluation and monitoring, can support trust-building and reduce the operational risks associated with premature autonomy.

In addition, organizations should invest in governance capabilities alongside technical infrastructure. The findings indicate that successful scaling requires internal management structures capable of managing probabilistic system behavior, defining acceptable risk thresholds, and ensuring compliance with regulatory and ethical standards. Without such governance arrangements, even technically successful pilots are unlikely to progress to sustained operational use.

For ventures operating within platform ecosystems, the study highlights the importance of actively managing dependency on dominant infrastructure providers. While reliance on established platforms may be unavoidable in the short term, entrepreneurs and adopters should seek to preserve strategic flexibility through modular system design, interoperability planning, and continuous monitoring of platform governance changes. Such practices can help mitigate the risks associated with the infrastructural power held by external suppliers while still maintaining access to critical technological resources.

The recommendations emphasize that the successful scaling of AI agent frameworks depends less on achieving maximal autonomy and more on developing organizational and strategic capabilities that allow agents to be deployed responsibly, reliably, and at scale. By aligning technical development with data governance, human oversight, and ecosystem awareness, organizations can move beyond experimental adoption and toward sustainable business integration of agentic AI.

BIBLIOGRAPHY AND REFERENCES

1. Acemoglu, D. (2024). The simple macroeconomics of AI. Economic Policy, 40, 13-58.
2. Adekunle, O. K., Ologbon, E. G., & Abiodun, S. J. (2024). Artificial Intelligence (AI) Adoption and Startup Success Rates. International Journal of Research and Innovation in Social Science, 8(11), 10235.
3. Black, J., Dascalu, D., Hughes, M., Wilkinson, B., Ryan, M., Bregman, A., Carlyon, P., Cheung, J., Freedman, L., Lucas, R., Patalano, A., Porter, P., Quimbre, F., Stockwell, S., & Virdee, M. (2024). Strategic competition in the age of AI: Emerging risks and opportunities from military use of artificial intelligence (Report No. RRA3295-1). RAND Corporation.
4. Caloffi, A., Colovic, A., Rizzoli, V., & Rossi, F. (2023). Innovation intermediaries' types and functions: A computational analysis of the literature. Technological Forecasting & Social Change, 189, 122351.
5. Canboy, B., & Khlif, W. (2025). Beyond efficiency: Revisiting AI platforms, servitization and power relations from a critical perspective. International Journal of Production Economics, 282, 109550.
6. Colovic, A., Caloffi, A., Rossi, F., Paladini, S., & Bagherzadeh, M. (2024). Innovation intermediaries and emerging digital technologies. Technovation, 133, 103022.
7. Engert, M., Hein, A., & Krcmar, H. (2025). Self-organization and governance in digital platform ecosystems: An information ecology approach. MIS Quarterly, 49(1), 91-122.
8. Foster, C. (2024). Openness in AI and downstream governance: A global value chain approach. Competition & Change.
9. Garro, A., Mühlhäuser, M., Tundis, A., Baldoni, M., Baroglio, C., Bergenti, F., & Torroni, P. (2025). Intelligent agents: Multi-agent systems. In Encyclopedia of Bioinformatics and Computational Biology. Elsevier.
10. Geng, S., & Liu, S. (2025). An agent-based framework for resilience analysis of service networks. Reliability Engineering and System Safety, 253, 110523.
11. Ghazawneh, A., & Henfridsson, O. (2013). Balancing platform control and external contribution in third-party development: The boundary resources model. Information Systems Journal, 23(2), 173-192.
12. Google Cloud. (2025). The ROI of AI 2025: How agents are unlocking the next wave of AI-driven business value.
13. Hadfield, G. K., & Koh, A. J. (2025). An Economy of AI Agents [Working paper]. Johns Hopkins Department of Computer Science; MIT Department of Economics.

14. Hafeez, S., Shahzad, K., & De Silva, M. (2025a). Enhancing digital transformation in SMEs: The dynamic capabilities of innovation intermediaries within ecosystems. Long Range Planning, 58, 102525.
15. Hafeez, S., Shahzad, K., Helo, P., & Mubarak, M. F. (2025b). Knowledge management and SMEs' digital transformation: A systematic literature review and future research agenda. Journal of Innovation & Knowledge, 10, 100728.
16. Hang, H., & Chen, Z. (2022). How to realize the full potentials of artificial intelligence (AI) in digital economy? A literature review. Journal of Digital Economy, 1, 180-191.
17. Hein, A., Schreieck, M., Riasanow, T., Setzke, D. S., Wiesche, M., Böhm, M., & Krcmar, H. (2020). Digital platform ecosystems. Electronic Markets, 30, 87-98.
18. Holland, C., McCarthy, A., Ferri, P., & Shapira, P. (2024). Innovation intermediaries at the convergence of digital technologies, sustainability, and governance: A case study of AI-enabled engineering biology. Technovation, 129, 102875.
19. Irman, A., et al. (2025). Artificial Intelligence (AI) adoption in business: Opportunities and challenges. International Journal of Business, Economics and Social Development, 6(1), 99-104.
20. Jobstreibizer, J., Beliaeva, T., Ferasso, M., Kraus, S., & Kallmuenzer, A. (2025). The impact of artificial intelligence on business models: A bibliometric-systematic literature review. Management Decision, 63(13), 372-396.
21. Juntunen, M. (2017). Business model change as a dynamic capability [Doctoral dissertation, University of Oulu]. Acta Universitatis Ouluensis G Oeconomica 94.
22. Kemell, K.-K., Wang, X., Nguyen-Duc, A., Grendus, J., Tuunanen, T., & Abrahamsson, P. (2020). 100+ metrics for software startups - common practices of using metrics. In Fundamentals of Software Startups: Essential Engineering and Business Aspects. Springer.
23. Kilian, K. A. (2025). Beyond accidents and misuse: Decoding the structural risk dynamics of artificial intelligence. AI & Society.
24. Kotu, V., Murphy, R. M., Solis, B., & Zilbershot, D. (2025). Enterprise AI Maturity Index 2025. ServiceNow.
25. Lari, K., Cant, K., & Evins, R. (2025). An agent-based framework for prioritizing building retrofits. Journal of Building Engineering, 107, 112661.
26. Luitse, D. (2024). Platform power in AI: The evolution of cloud infrastructures in the political economy of artificial intelligence. Internet Policy Review, 13(2).
27. Masiero, S., Qosaja, J., & Cutrona, V. (2024). Digital datasheet model: Enhancing value of AI digital platforms. Procedia Computer Science, 232, 149-158.
28. McKinsey & Company. (2024). The state of AI in 2024: Generative AI's breakout year. McKinsey Global Survey.

29. Meyer, J.-J. C. (2014). Logics for intelligent agents and multi-agent systems. In D. M. Gabbay & J. Woods (Eds.), Handbook of the History of Logic: Vol. 9. Computational Logic (pp. 629-658). Elsevier.
30. Osterwalder, A., & Pigneur, Y. (2010). Business model generation: A handbook for visionaries, game changers, and challengers. John Wiley & Sons.
31. Oželienė, D. (2024). Methodological guidelines for preparation, defence and evaluation of Master's theses (project). Vilnius University Business School.
32. Parker, G. G., Van Alstyne, M. W., & Choudary, S. P. (2016). Platform revolution: How networked markets are transforming the economy - and how to make them work for you. W. W. Norton & Company.
33. Peasley, D., Kuplicki, R., Sen, S., & Paulus, M. (2025). Leveraging large language models and agent-based systems for scientific data analysis: Validation study. JMIR Mental Health, 12, e68135.
34. Russell, S., & Norvig, P. (2021). Artificial intelligence: A modern approach (4th ed.). Pearson.
35. Sayyed-Alikhani, A., Chica, M., & Mohammadi, A. (2021). An agent-based system for modeling users' acquisition and retention in startup apps. Expert Systems With Applications, 176, 114861.
36. Sjödin, D., Parida, V., & Kohtamäki, M. (2023). Artificial intelligence enabling circular business model innovation in digital servitization: Conceptualizing dynamic capabilities, AI capacities, business models and effects. Technological Forecasting & Social Change, 197, 122903.
37. Sohn, S. M. (2024). The three levels of artificial intelligence (AI) adoption framework and the six levels of autonomous companies.
38. Tang, Y., Chen, K., Yue, L., Fan, J., Zhou, C., Li, X., ... & Zhang, M. (2025). Empowering real-world: A survey on the technology, practice, and evaluation of LLM-driven industry agents. arXiv preprint.
39. Teece, D. J. (2007). Explicating dynamic capabilities: The nature and microfoundations of (sustainable) enterprise performance. Strategic Management Journal, 28(13), 1319-1350.
40. Teece, D. J. (2010). Business models, business strategy and innovation. Long Range Planning, 43(2-3), 172-194.
41. Tupe, V., & Thube, S. (2025). AI agentic workflows and enterprise APIs: Adapting API architectures for the age of AI agents.
42. Wang, Y., et al. (2023). Differentially private consensus and distributed optimization in multi-agent systems: A review [Preprint]. Submitted to Neurocomputing.

43. Warner, K. S. R., & Wäger, M. (2019). Building dynamic capabilities for digital transformation: An ongoing process of strategic renewal. Long Range Planning, 52(3), 326-349.
44. Wei, R., & Pardo, C. (2022). Artificial intelligence and SMEs: How can B2B SMEs leverage AI platforms to integrate AI technologies? Industrial Marketing Management, 107, 466-483.
45. Xu, L., Almahri, S., Mak, S., & Brintrup, A. (2024). Multi-agent systems and foundation models enable autonomous supply chains: Opportunities and challenges. IFAC PapersOnLine, 58(19), 795-800.
46. Yehudai, A., Eden, L., Li, A., Uziel, G., Zhao, Y., Bar-Haim, R., Cohan, A., & Shmueli-Scheuer, M. (2025). Survey on evaluation of LLM-based agents. arXiv preprint.
47. Yu, S., & Sekiguchi, T. (2024). Platform-dependent entrepreneurship: A systematic review. Administrative Sciences, 14(12), 326.
48. Zhan, Y., Xiong, Y., Han, R., Lam, H. K. S., & Blome, C. (2024). The impact of artificial intelligence adoption for business-to-business marketing on shareholder reaction: A social actor perspective. International Journal of Information Management, 76, 102768.
49. Zhao, B., Foo, L. G., Hu, P., Theobalt, C., Rahmani, H., & Liu, J. (2025). LLM-based agentic reasoning frameworks: A survey from methods to scenarios. Journal of the ACM, 37(4), Article 111.
50. Ziakis, C., & Kavoura, A. (2025). Signals of growth: Comparing web metrics performance in B2B and B2C Software-as-a-Service startups. University of Macedonia & University of West Attica.

ANNEXES

Annex 1

Table 2. *Overview of Expert Interviews*

Interview Code	Ecosystem Role	Industry	Area of Expertise	Interview Duration
I1	Creator	AI infrastructure provider	Agent architecture, system design	21min
I2	Integrator	Engineering services	AI deployment, system integration	27min
I3	Adopter	FinTech startup	Porduct strategy, compliance	41min
I4	Adopter	E-commerce startup	Operations, automation	25min
I5	Adopter	Martech startup	Data-drive marketing systems	25min

Source: compiled by the author

Annex 2

Semi-structured interview guides

1. Interview guide - Creators

Background and role:

1. Could you briefly describe your current role and responsibilities related to AI or agent-based systems?
2. How long have you been working with AI, automation, or agent-based technologies?

Main questions:

3. Which types of customers or industries do you typically support in AI-related projects?

4. What parts of the solution are you most involved in (e.g., architecture, APIs, tooling, integration patterns)?

5. Could you describe one recent enterprise workflow or use case you supported where AI or automation played a significant role?

6. What tools, systems, or APIs were involved in that workflow?

7. In that workflow, how does the AI system or agent interact with the client's existing tools or applications?

8. What factors make integration with enterprise systems easier or more difficult (e.g., APIs, data quality, permissions, orchestration)?

9. Before moving a system like this into production, what minimum safeguards or checks must be in place?

10. How do you help customers balance agent autonomy with safety and human oversight?

11. Across your projects, what are the most common barriers that slow down or prevent adoption?

12. How do customers typically express or prioritize concerns related to accuracy, integration cost, privacy, security, or ownership?

13. Which delivery model works best for enterprise workflows (e.g., self-service API, packaged SaaS solution, managed service)?

14. What factors influence the choice of delivery model?

15. If advising a customer to run a small 4–6-week pilot, what would be the essential steps?

16. What single KPI or outcome best indicates whether a pilot is successful?

17. Is there anything important we did not discuss that you believe matters for adopting or scaling agent-based AI systems?

2. **Interview guide - Integrators**

Background and role:

1. Could you briefly describe your current role and responsibilities in AI or automation projects?

2. How long have you been working with AI, system integration, or digital transformation initiatives?

Main questions:

3. Which types of clients or industries do you primarily support in AI or automation projects?

4. At which stages of AI projects are you most involved (e.g., discovery, design, integration, deployment, support)?

5. Could you describe one recent workflow your team automated or enhanced using AI?

6. Which tools, data sources, and systems needed to work together in that project?

7. What usually makes integration with client systems straightforward, and what makes it difficult?

8. How do you bridge gaps between what AI systems require and what client systems can provide?

9. In the workflows you have automated, which parts can be executed autonomously and which require a human in the loop?

10. How do you determine the appropriate level of autonomy for each step in a workflow?

11. What governance or safety requirements do clients typically require before deployment?

12. Which governance or compliance constraints most often slow down implementation?

13. What internal capabilities or team structures make organizations more successful in adopting AI-driven workflows?

14. Where do you most often observe organizational readiness gaps (e.g., skills, processes, data, leadership, risk tolerance)?

15. Across projects, what are the biggest barriers that slow down or block adoption?

16. Are these barriers more commonly technical, organizational, or governance-related?

3. Interview guide - Adopters

Background and role:

1. Could you briefly describe your role and responsibilities within your organization?

2. How are you involved in decisions related to AI, automation, or digital transformation?

Main questions:

3. How does AI or automation fit into your organization’s current priorities?

4. Could you describe one workflow or process where your organization recently explored or implemented AI or automation?

5. What systems, tools, or teams are involved in that workflow?

6. Which steps in the workflow are currently the most painful, slow, or repetitive?

7. Where do you see opportunities for more autonomous, multi-step AI-driven execution across tools?

8. When experimenting with AI or automation, what makes integration into your existing systems easier or harder?

9. How important is interoperability with your current tools?

10. Before using AI to automate real work, what safeguards or checks must be in place in your organization?

11. How do you think about responsibility and ownership when AI systems make decisions or take actions?

12. What internal capabilities or teams are essential for adopting AI systems that automate multi-step work?

13. What internal factors slow down adoption (e.g., skills, processes, risk tolerance, data readiness)?

14. How do you assess whether an AI-driven workflow delivers value, and which metrics matter most?

15. What makes it difficult to justify investment in agent-based AI solutions?

16. Which delivery model best fits your organization (e.g., self-service API, packaged SaaS solution, managed service)?

17. Why does this delivery model work best for your organization?

18. If running a 4–6-week pilot, how would you scope it?

19. What would indicate that the pilot is successful and worth scaling?

20. What would prevent your organization from moving beyond the pilot stage?

21. Is there anything important we did not discuss that you believe matters for adopting or scaling agent-based AI systems?

Annex 3

Table 3. *Methodological process overview*

Research Stage	**Description**
Research design	Qualitative, interpretivist, exploratory
Sampling strategy	Purposive expert sampling
Data collection	Semi-structured online interviews
Data preparation	AI-assisted transcription with manual verification
Analysis method	Reflexive Thematic Analysis
Validation measures	Role-based perspective triangulation, reflexivity

Source: compiled by the author

Annex 4

Table 4*. Overview of empirical themes*

Theme	**Roles Contributing**	**Empirical Focus**

Data readiness and quality	Creator, Integrator, Adopters	Data structure, preparation effort
Human oversight	Integrator, Adopters	Trust, regulation, evaluation
Integration barriers	Integrator, Adopters	APIs, legacy systems
Governance and security	Adopters	Compliance, data segmentation
Strategic value	Integrator, Adopters	ROI, pilot evaluation

Source: compiled by the author