VILNIUS UNIVERSITY

FACULTY OF ECONOMICS AND BUSINESS ADMINISTRATION

GLOBAL BUSINESS AND ECONOMICS

Jelena Burbė

MASTER THESIS

KONKURENCINIO PRANAŠUMO KŪRIMAS GAMYBOJE: ĮMONIŲ ARCHITEKTŪROS BRANDOS VAIDMUO REALIZUOJANT DIRBTINIO INTELEKTO VERTĘ - PRIEŽASTINĖS ANALIZĖS TYRIMAS	**ARCHITECTING COMPETITIVE ADVANTAGE IN MANUFACTURING: THE ROLE OF ENTERPRISE ARCHITECTURE MATURITY IN REALIZING AI VALUE - A CAUSAL INFERENCE ANALYSIS**

Master degree student ______________________
(signature)

Supervisor______________________
(signature)

Supervisor Prof. Dr. Martina Dal Molin

Date of submission of Master Thesis:
Ref. No.

Vilnius, 2026

CONTENTS

INTRODUCTION

Manufacturing firms increasingly deploy Artificial Intelligence (AI) to improve operational performance - reducing downtime, stabilizing quality, optimizing yield, and accelerating decision cycles, yet the realized value of these initiatives is frequently uneven across plants and over time. In many organizations, AI delivers convincing results at the pilot stage but does not scale into a stable, enterprise-wide capability, resulting in volatile outcomes and persistent uncertainty regarding whether expected returns will materialize. This thesis investigates this value-realization problem in manufacturing by examining whether Enterprise Architecture (EA) maturity helps explain why some firms repeatedly convert AI initiatives into scalable and persistent operational improvements, while others remain constrained by fragmented, local successes.

Relevance of the topic. From a practical perspective, the topic is practically relevant because manufacturing organizations typically operate in highly complex IT/OT landscapes where data, processes, and decision rights are distributed across heterogeneous systems. In such environments, the cost of coordination and integration is a binding constraint: even technically successful AI solutions may fail to deliver sustained value if they rely on unstable data pipelines, inconsistent semantics, or non-standardized execution processes. Practice-oriented evidence supports this relevance. Manufacturing digital transformation programs often struggle to scale precisely because cross-layer alignment between business processes, organization, and technology is missing, which increases the probability of “pilot traps” (Behrendt et al., 2021). In parallel, large-scale EA industry studies report that IT complexity has increased for many firms while mature EA programs remain scarce, suggesting a structural governance gap that can undermine AI execution at scale (LeanIX, 2019; Bizzdesign, 2025).

The topic is theoretically relevant because existing research offers strong, but partially disconnected explanations of digital value creation. EA research emphasizes standardization, integration, and governance as a “foundation for execution”, enabling repeatability and scalability of initiatives across organizational units (Ross, Weill, & Robertson, 2006). AI and digital transformation research similarly argue that value from AI depends on complementary investments in operating models and organizational learning rather than on technology deployment alone (Agrawal, Gans, & Goldfarb, 2018; Iansiti & Lakhani, 2020). Meanwhile, empirical IT capability research shows that firm-level IT capabilities are associated with improved performance, reinforcing the view that organizational capabilities, not tools, drive durable advantages (Bharadwaj, 2000; Brynjolfsson, Hitt, & Kim, 2011). However, despite this conceptual complementarity, the intersection remains insufficiently tested empirically in manufacturing settings characterized by extreme integration complexity and heterogeneous plant architectures.

Level of exploration of the topic. EA research has mapped a broad set of EA topics and benefits, but measurement approaches and empirical outcomes remain uneven across the field (Kotusev, 2017). Empirical work suggests that EA-induced capabilities and architectural insight can influence organizational and project performance, indicating that EA is not merely descriptive, but capability-building (Foorthuis et al., 2016; Niemi & Pekkola, 2020). Manufacturing-oriented empirical studies indicate that EA strategic orientation and assimilation act as capabilities that support agility and performance, consistent with the view, that architecture is competence rather than a documentation practice (Hazen et al., 2017). The digital transformation literature contains strong evidence that many organizations fail to scale industrial AI/IoT initiatives due to operating model and architecture misalignment, despite local pilot success (Behrendt et al., 2021). Finally, methodological advances in causal inference, especially in high-dimensional settings, provide frameworks such as double/debiased machine learning to estimate structural parameters when confounding is complex and traditional models are fragile (Chernozhukov et al., 2018; Athey & Imbens, 2019). Taking them together, these streams imply that architectural capability should matter for AI value realization, but they stop short of providing a manufacturing-focused empirical account of scaling, persistence, and variance in AI outcomes as functions of EA maturity.

Against this background, the research gaps addressed in this thesis are substantive. First, while the "foundation for execution" logic is well established, there is limited empirical evidence linking EA maturity to manufacturing AI outcomes in a way that is comparable across firms and time (Ross et al., 2006; Kotusev, 2017). Second, the pilot-to-scale problem is well documented in industrial transformation evidence, yet the role of EA maturity as a structural determinant of scaling remains insufficiently tested in a firm-year setting (Behrendt et al., 2021; LeanIX, 2019). Third, persistence of AI-driven performance gains, whether improvements endure and accumulate, is frequently discussed as a transformation challenge, but is rarely operationalized as a testable outcome linked to architecture maturity (Niemi & Pekkola, 2020; Hazen et al., 2017). Fourth, outcome variance, why AI results are stable for some firms and volatile for others, remains underexplained in empirical research, despite strong industry evidence that maturity and governance differences shape predictability of outcomes (Bizzdesign, 2025; LeanIX, 2019). These gaps are particularly important in manufacturing, where architectural fragmentation can turn AI into a high-risk investment: not because models fail technically, but because organizations cannot reliably execute and scale what works locally.

Literature gap. Although EA maturity is widely theorized as a "foundation for execution", existing research rarely links EA maturity to realized operational AI value in a way that is comparable across firms and time, especially in manufacturing contexts with high IT/OT integration constraints. In parallel, the pilot-to-scale problem is frequently documented, yet the architectural conditions that enable scaling, persistence, and variance reduction remain under-tested empirically in firm-year designs. This thesis addresses this gap by testing whether EA maturity functions as a structural threshold for moving beyond pilots and stabilizing AI outcomes.

Novelty of the thesis. To address the substantive gap in the scaling of AI initiatives in manufacturing, EA maturity is treated in this study as a firm-specific organizational capability reflecting the degree of architectural standardization, integration, and governance. Moving beyond descriptive surveys of framework compliance, the research operationalizes EA maturity through a TOGAF[1]-aligned proxy index derived from observable and publicly verifiable indicators, enabling objective longitudinal and cross-firm comparison of the firm's architectural "foundation for execution".

Methodologically, the study adopts a causal identification logic consistent with the Double Machine Learning (DML) framework, allowing EA maturity to be examined as a structural conditioning factor while accounting for high-dimensional firm-level confounders such as legacy technical debt and prior digital maturity. The dependent phenomenon is the realized value of AI in manufacturing, captured through an operational value index and, where explicitly disclosed, financial ROI.

By focusing on realized outcomes rather than adoption counts or technical performance, the thesis provides a decision-relevant lens for evaluating whether an organization's architectural state functions as an enabling threshold - or a structural barrier - to scaling AI as a modern "prediction machine" in complex industrial environments.

Aim of the thesis. Starting from this background, the aim of this thesis is to evaluate how EA Maturity influences AI-driven value realization in manufacturing organizations, with emphasis on scaling beyond pilots, persistence of improvements over time, and variance in realized outcomes, and to examine whether higher EA maturity is associated with reduced divergence between expected and realized outcomes where expectation data are observable.

Objective of the thesis. The thesis addresses four research questions:

(1) How is EA maturity associated with the level of realized operational AI value across manufacturing firm-years?

(2) Does higher EA maturity increase the probability that AI initiatives transition from localized pilots to multi-site or enterprise-wide deployment?

[1] TOGAF (The Open Group Architecture Framework) is a widely used enterprise architecture framework that structures enterprise design into four core architectural domains: Business, Data, Application, and Technology Architecture.

(3) Are AI-related performance improvements more persistent over time in organizations with higher EA maturity than in organizations with fragmented architectures?

(4) For comparable AI investments where observable, do higher-EA-maturity organizations exhibit lower variance in realized AI-related operational outcomes?

The hypotheses tested in this study are formulated as follows:

H1 – Bias correction

Organizations with higher EA maturity exhibit smaller differences between ex-ante ROI expectations and ex-post realized outcomes of AI initiatives.

H2 – Pilot-to-scale transition

EA maturity positively increases the probability that AI initiatives transition from localized pilots to multi-site or enterprise-wide deployment.

H3 – Outcome persistence

Performance improvements associated with AI initiatives persist longer in organizations with higher EA maturity than in organizations with fragmented architectures.

H4 – Variance reduction

For comparable AI investments, organizations with higher EA maturity show lower variance in realized AI-related operational outcomes.

Methods deployed in the Master thesis. The study adopts a quantitative, non-experimental firm-year design based on publicly verifiable secondary sources. EA maturity is operationalized as a composite proxy aligned with TOGAF architectural domains, while AI outcomes are captured through a primary ordinal measure of realized operational AI value and, where explicitly disclosed, financial AI Return on Investment (AI ROI). Methodologically, the thesis is designed around a causal inference framework and draws on established econometric logic and modern causal methods described in the literature; however, the empirical implementation prioritizes transparent, hypothesis-consistent evidence grounded in observable patterns of scaling, persistence, and variance under manufacturing constraints.

Description of structure of the thesis. The thesis is structured as follows. The first chapter develops the theoretical foundation linking EA maturity, AI as a general-purpose technology, and manufacturing complexity. The second chapter presents empirical methodology, including variable operationalization and identification logic. The third chapter reports empirical results and strategic implications, linking findings back to the theoretical framework. The thesis concludes with conclusions and recommendations for practice and future research.

In accordance with the Vilnius University Faculty of Economics and Business Administration methodological guidelines, this thesis discloses the use of AI tools during preparation and clarifies the scope of their use. Microsoft Copilot and Google NotebookLM were used to streamline literature gathering workflows, support thematic organization of sources, and assist with linguistic refinement (grammar and clarity). AI-assisted outputs were carefully reviewed and iteratively revised; sources and claims were judged against original materials, definitions were checked for consistency, data handling was debugged, and the final text was refined to academic standard. The author remains fully responsible for the research design, empirical coding, interpretation of results, and academic integrity of the thesis.

1. THEORETICAL BASIS FOR ENTERPRISE ARCHITECTURE MATURITY AND AI VALUE

1.1. Conceptualizing Enterprise Architecture as a strategic foundation for firm execution

1.1.1. The role of business process standardization and integration

In EA research, a central distinction is drawn between initiatives that optimize local processes and those that contribute to a broader foundation for execution, enabling scalable and repeatable organizational performance. From a capability perspective, initiatives that are locally sensible often prioritize the immediate functional needs of a single business unit, frequently under time pressure or operational strain, while leaving enterprise-wide process standardization and system integration unaddressed (Ross, Weill, & Robertson, 2006; Kotusev, 2017). Although such initiatives may deliver rapid local benefits, they tend to produce islands of automation: isolated systems connected through complex and fragile integrations, which increase long-term coordination costs and architectural fragility (Ross et al., 2006).

By contrast, initiatives that strengthen a foundation for execution are deliberately designed around shared business process definitions, common data semantics, and a single authoritative representation of core operational entities. Rather than optimizing individual functions in isolation, these initiatives emphasize standardization and integration as organizational capabilities, allowing the firm to scale operations, enter new markets, and introduce new digital capabilities with significantly lower marginal complexity (Ross et al., 2006; Foorthuis et al., 2016). The strategic value of this approach becomes particularly visible when competitors struggle with hidden interdependencies that were never explicitly mapped, governed, or standardized, constraining their ability to adapt or scale.

Importantly, adopting a foundation-for-execution perspective does not imply rejecting all locally initiated solutions. Instead, it requires evaluating local decisions through a system-wide perspective, ensuring that corrective or intermediate solutions represent deliberate steps toward greater standardization rather than reinforcing fragmentation. Capability-oriented EA research emphasizes that architectural maturity emerges through cumulative choices aligned with enterprise-level principles, not through isolated design decisions (Niemi & Pekkola, 2020). This shift reflects a move from departmental optimization toward end-to-end process thinking, where long-term organizational coherence outweighs short-term local convenience.

Business process standardization further enables cross-unit communication through a shared operational language, making it possible to define, compare, and govern performance consistently across the organization. In practice, this involves structuring key performance indicators around agreed definitions, identifying authoritative data sources, formalizing calculation logic and ownership, and linking performance metrics to decision rules and business consequences (Ross et al., 2006). Indicators such as requirement clarity, scope stability, rework resulting from specification gaps, architectural coupling, legacy dependency, and delivery predictability act as feedback mechanisms that expose structural weaknesses and guide continuous improvement.

Through these mechanisms, EA evolves beyond an IT concern into a strategic foundation for organizational learning. By systematically surfacing the hidden costs of silos and making architectural trade-offs explicit to decision-makers, the organization develops a more disciplined and mature response to business demands. The outcome is not merely better systems, but an enterprise-level capability to align local initiatives with global strategy, ensuring that incremental decisions consistently reinforce, rather than erode the organization's foundation for execution (Ross et al., 2006; Foorthuis et al., 2016).

1.1.2. Evolution of organizational learning through Architecture Maturity stages

In EA maturity models, the transition toward an optimized core represents a critical inflection point in organizational transformation. Research on EA as a strategic foundation for business execution emphasizes that this stage requires the institutionalization of enterprise-wide governance over core operational processes, rather than incremental technical alignment alone (Ross, Weill & Robertson, 2006). While earlier stages, such as siloed or partially standardized architectures, typically focus on hardware consolidation or basic process harmonization, the optimized core demands that discretionary control over key operational processes be replaced by shared, cross-organization standards and decision rights. This shift is frequently experienced as more challenging than technical standardization because it disrupts established local authority structures and deeply ingrained ways of working.

At this stage, managers are no longer evaluated primarily on localized performance optimization but are expected to operate within rigid, cross-plant process models designed to support global efficiency. Enterprise-level execution research highlights that such transitions often generate resistance precisely because they redefine accountability, constrain managerial autonomy, and reveal previously implicit trade-offs between local flexibility and enterprise coherence (Ross et al., 2006). As a result, the optimized core stage is commonly associated with heightened organizational tension rather than immediate and visible performance gains.

The transition toward an optimized core frequently produces a period of disorientation among operational leaders, as standardized and data-driven workflows impose cognitive and organizational demands that exceed prior experience with locally adapted processes. Once an organization moves decisively into this phase, reversibility becomes extremely limited. Process integration reaches a level where ad hoc rollbacks or panic-driven deviations threaten systemic stability rather than correcting isolated deficiencies. In such contexts, leadership attention must shift from technical implementation toward actively managing resistance, reinforcing governance mechanisms, and sustaining transformation momentum. Failure to do so risks enterprise-level breakdown rather than localized inefficiency (Ross et al., 2006).

At the same time, architectural maturity within manufacturing organizations is rarely uniform across all systems or domains. Different process areas and platforms may coexist at varying levels of maturity; a condition widely acknowledged in EA literature. For this reason, modular architectural design plays a critical role in enabling progress without imposing premature uniformity. Modularization allows for reuse and standardization where enterprise value is highest, while still permitting controlled local adaptations when justified by operational realities. However, this flexibility must be balanced against the well-documented negative effects of persistent silos on end-to-end process performance and decision quality.

By continuously making architectural trade-offs explicitly linking design decisions to tangible business outcomes such as scalability, resilience, cost transparency, and risk exposure - organizations progressively evolve not only in technical capability but also in organizational learning maturity. EA thus functions as a discipline mechanism that supports more context-aware and system-oriented decision-making, reducing the tendency to revert to fragmented, locally optimized solutions under operational pressure. In this sense, maturity at the optimized core stage reflects an organization's capacity to consistently privilege long-term execution capability over short-term convenience, reinforcing rather than undermining its foundation for execution (Ross et al., 2006).

1.2. Economic implications of Artificial Intelligence as a general-purpose technology

1.2.1. The impact of falling prediction costs on strategic decision-making

Recent economic analyses conceptualize AI as a technology that fundamentally lowers the cost of prediction across a wide range of decision contexts. In manufacturing, declining prediction costs, enabled by machine learning and real-time analytics, allow firms to detect product defects, process deviations, and demand fluctuations earlier and with greater precision. Consistent with the "cheap prediction" framework, reductions in prediction costs alter the relative economics of decision-making uncertainty becomes cheaper to resolve, while delayed or incorrect action becomes comparatively more expensive (Agrawal, Gans, & Goldfarb, 2018). As a result, the strategic bottleneck shifts from forecasting what will happen to determining which predicted futures warrant action, elevating the role of judgment and organizational design (Agrawal et al., 2018).

Framed as a general-purpose technology, AI resembles earlier GPTs in that its economic impact depends on complementarities and system-wide co-invention rather than isolated point improvements (Bresnahan & Trajtenberg, 1995). Prediction capability alone does not generate durable competitive advantage unless firms redesign processes, decision rights, and integration mechanisms around it. This property of AI strengthens the importance of organizational and architectural alignment and directly anticipates the role of human judgment as a complementary asset, which is examined in the subsequent subsection.

From an economic perspective, the marginal informational value of additional data exhibits diminishing returns. In statistical terms, improvements in estimation precision typically scale sub-linearly with sample size, approximately proportional to $1/\sqrt{n}$, implying that once baseline uncertainty has been reduced, progressively larger increases in data volume are required to achieve equivalent gains in predictive accuracy (NIST/SEMATECH, 2022). Consequently, as competitors converge on comparable sensing infrastructures and data platforms, data accumulation alone becomes a weaker differentiator unless paired with superior algorithms and high-quality judgment that translate predictions into actionable decisions (Agrawal et al., 2018).

This shift in relative costs makes investment-intensive strategies, such as anticipatory shipping, condition-based maintenance, and outcome-based service contracts - economically viable for traditional manufacturers. Cheaper prediction supports earlier intervention and more precise pricing of operational risk, enabling firms to internalize uncertainty that was previously too costly to manage (Agrawal et al., 2018). Rather than relying solely on transfer-of-ownership models, manufacturers can increasingly offer guaranteed performance or uptime, transforming uncertainty into a monetizable dimension of value creation. In parallel, decentralized intelligence embedded within production and logistics systems allows deviations to be corrected before defects propagate downstream, shifting enterprise systems from post-hoc correction toward continuous, prediction-driven orchestration across production, inventory, and service processes (Agrawal et al., 2018).

Finally, diminishing returns to indiscriminate data accumulation necessitate a strategic rebalancing from a “data-first” posture - collecting as much data as possible - to a “business-first” posture that identifies the specific uncertainty that matters, defines the performance metric that captures it, and generates the minimum data required to resolve it. This transition is not merely a governance preference but represents the economically efficient response in environments where prediction is cheap and data abundance no longer guarantees advantage (Agrawal et al., 2018; Bresnahan & Trajtenberg, 1995). Under these conditions, judgment becomes the scarce resource that determines which predictions enter the decision space and how they are translated into action, motivating the analysis in the following subsection.

This complements AI economics arguments that prediction becomes cheaper (Agrawal et al., 2018) but value depends on organizational redesign and operating-model complements (Iansiti & Lakhani, 2020), and it is consistent with evidence that practical diffusion and task redesign constraints slow the realization of AI gains beyond technical feasibility (Brynjolfsson & Mitchell, 2017).

1.2.2. Human judgment as a complementary asset to machine intelligence

Economic theories of AI emphasize that as prediction becomes inexpensive and widely accessible, the scarce resource shifts to judgment: the human capacity to assign value to outcomes and decide which predicted futures warrant action (Agrawal, Gans, & Goldfarb, 2018). In manufacturing contexts, this distinction is particularly important because statistically accurate predictions may nonetheless be operationally infeasible. As a result, AI-generated outputs should not be treated as decisions, but rather as inputs that must pass feasibility filters before entering the decision space - filters grounded in enterprise system logic, physical constraints, safety requirements, and execution capability (Agrawal et al., 2018).

For example, an AI model may identify a statistically "optimal" warehouse picking route, yet that solution may be infeasible once floor-load limits, aisle geometry, equipment turning radii, safety zoning, or regulatory constraints are considered. In such cases, numerical optimality alone is insufficient. Judgment is required to evaluate trade-offs under real operational constraints and to determine whether intervention is permissible, controllable, and safe (Agrawal et al., 2018). This distinction aligns with the broader "prediction versus judgment" framework, which separates knowing the likely state of the world from determining the value and feasibility of actions within that state.

Beyond localized operational decisions, this filtering role of judgment directly extends to EA choices. Architectural decisions, such as separating snapshot-based analytical data from transactional execution data, or favoring modular and reversible designs over tightly coupled automation, determine whether prediction remains a flexible input into human judgment or becomes prematurely embedded in execution logic. From a general-purpose technology perspective, AI generates value not through prediction accuracy alone, but through complementarities with organizational design, governance, and learning processes (Bresnahan &

Trajtenberg, 1995). Poor architectural choices can therefore convert a potentially valuable prediction capability into rigid automation, that constrains rather than enhances decision quality.

The principle often summarized as “don’t automate - obliterate” further clarifies why judgment must govern automation boundaries. Embedding prediction into poorly designed or misaligned processes tends to amplify inefficiency rather than eliminating it. Business process redesign must precede automation so that prediction reduces uncertainty in a rearchitected value stream instead of accelerating legacy dysfunction (Hammer, 1990). In manufacturing environments, sustainable competitive advantage thus arises not from superior prediction capability alone, but from the disciplined integration of low-cost prediction with domain-informed judgment. Only when analytically sound predictions pass feasibility filters and are translated into executable decisions can AI contribute to operationally viable and strategically sustainable outcomes (Agrawal et al., 2018).

1.3. Structural dynamics of modularity and complexity in manufacturing environments

1.3.1. Modular operators and the decentralization of design options

Research on modularity emphasizes that sustained adaptability in complex systems depends on the decomposition of functionality into loosely coupled components governed by stable interfaces. In manufacturing information systems, this logic is particularly important for preventing new functionality from becoming permanently embedded in legacy ERP landscapes. Modularity theory identifies a set of modular operators including splitting, substituting, and augmenting, that allow system designers to restructure functionality without destabilizing the system as a whole (Baldwin & Clark, 2000). When applied consistently, these operators enable evolution of business logic independently of specific technical implementations.

From this perspective, core business logic is encapsulated behind stable interfaces, while implementation details and technical complexity are deliberately hidden from the rest of the system. Whether functionality is realized through stored procedures, standardized modules, services, or future AI-based components is secondary to the architectural principle that design decisions remain reversible. Clean Architecture patterns operationalize this separation by enforcing clear boundaries between domain logic and infrastructure concerns, reducing the risk that experimental or rapidly evolving solutions harden prematurely into long-term technical debt.

This architectural approach allows functionality to evolve incrementally. Modules can be substituted or augmented without triggering cascading changes across the ERP core, and responsibilities can be split as complexity or change frequency increases. By treating modules as replaceable units rather than embedding logic directly into transactional execution flows, design options remain decentralized and reversible. Modularity thus preserves option value: the ability to adapt, recombine, or retire components as business conditions and technological capabilities evolve (Baldwin & Clark, 2000).

An additional benefit of modularization is fault isolation. When a component behaves unexpectedly, its impact can be contained locally, enabling targeted diagnostic and root-cause analysis rather than system-wide investigation. This containment reduces operational risk and limits the propagation of failure across tightly coupled processes. In this sense, modular architecture functions not only as a development strategy but also as an operational risk-management mechanism within complex manufacturing environments.

More broadly, modularity mediates the relationship between architectural complexity and organizational learning. By preventing tight coupling between prediction, execution, and infrastructure layers, modular design ensures that emerging capabilities, such as AI-driven analytics, can be explored, evaluated, and refined without prematurely constraining future choices. This aligns with general-purpose technology theory, which emphasizes that value emerges from system-level recombination and complementarities rather than from fixed technical solutions (Bresnahan & Trajtenberg, 1995). Through disciplined application of modular operators, EA mitigates architectural hardening and sustains the decentralization of design options necessary for long-term adaptability.

1.3.2. Rearchitecting the firm to overcome information silos and technical debt

EA research highlights that architectural complexity does not necessarily prevent automation but increasingly determines how automation fails. In environments characterized by implicit, unmapped integrations and accumulated technical debt, automation can proceed silently while eroding functional correctness beneath the surface. Rather than failing visibly, such systems permit structurally flawed changes to appear technically successful, masking deeper semantic or architectural inconsistencies. This condition reflects a form of architectural opacity in which system behavior becomes decoupled from business intent.

As integration complexity accumulates and dependencies remain implicit, automation ceases to fail loudly. Instead, systems enter a state in which changes are executed without explicit technical errors, even as downstream business functionality degrades. Architectural deficiencies remain undetected during execution because neither tooling nor system feedback surfaces semantic violations or structural mismatches. Consequently, defects become visible only during holistic system-level testing or, more critically, through discrepancies between expected business outcomes and observed system behavior, such as missing functionality, inconsistent data usage, or unexplained deviations in process execution.

This form of silent failure represents a particularly dangerous manifestation of technical debt. When causal links between architectural change, system behavior, and business impact cannot be traced, engineers and managers lose confidence in making further structural adjustments. Change avoidance emerges not from a lack of automation capability, but from the inability to reason reliably about how interventions propagate through the system. Under such conditions, architectural intervention is triggered not by system alerts or error messages, but by behavioral anomalies that require human interpretation rather than automated detection.

From a competitive-advantage perspective, architectural opacity does more than slow innovation - it renders the return on automation and AI investments fundamentally unpredictable. When organizations cannot causally connect architectural decisions to operational outcomes, they cannot estimate risk, benefit, or payback with confidence. This uncertainty undermines strategic decision-making and leads to deferral or abandonment of further automation initiatives, effectively stalling digital evolution despite the availability of advanced technologies.

Overcoming this condition requires deliberate rearchitecting efforts focused on reducing hidden integration paths, making dependencies explicit, and restoring modular boundaries. Clear architectural delineation enables both automation and human judgment to operate within well-defined decision scopes. In environments where automation produces no explicit error signals, judgment is activated by observed deviations in system behavior rather than by technical alerts alone, prompting architectural correction instead of further blind automation. This reinforces the argument developed in Section 1.2.2: as prediction and automation become cheaper and more powerful, human judgment shifts toward defining the boundaries, checkpoints, and semantic validations that determine where automation is permitted to act.

In manufacturing environments, unresolved architectural sprawl therefore converts system change into operational risk. Silent automation failures erode trust in data, planning, and execution systems, ultimately freezing further automation, AI adoption, and modernization at precisely the point where competitive advantage should be expanding. Rearchitecting the firm to restore transparency, modularity, and traceability is thus not merely a technical exercise, but a strategic prerequisite for harnessing advanced digital capabilities in a predictable and sustainable manner.

1.4. Behavioral dimensions of organizational decision-making under uncertainty

1.4.1. Cognitive biases and systematic errors in strategy formulation

Strategic decision-making under technological uncertainty is systematically influenced by cognitive biases, most notably the planning fallacy - the tendency to overestimate expected benefits while underestimating implementation complexity, time, and risk (Kahneman & Tversky, 1979; Kahneman, 2011). While individual optimism contributes to this bias, its persistence and amplification within large organizations are strongly shaped by the architectural context in which strategic decisions are formed.

In organizations with low EA maturity, architectural opacity constrains access to organizational memory artifacts, such as documented outcomes of prior ERP migrations, MES rollouts, cross-plant standardization initiatives, or earlier automation efforts that encountered integration constraints or data-model divergence. When such prior initiatives are not visible, comparable, or structurally documented, executives rely disproportionately on intuitive judgments about technological potential rather than on historically grounded assessments of feasibility, cost, and risk. Architectural opacity thus weakens the organization's ability to learn systematically from past outcomes.

This pattern aligns with established behavioral findings showing that decision makers often substitute complex forecasting problems with simpler narratives, particularly when systems are too complex to reason about analytically. Under conditions of high uncertainty and limited structural visibility, intuitive optimism becomes a cognitively efficient, but systematically biased, substitute for evidence-based evaluation (Kahneman, 2011). Architectural noise amplifies this effect by obscuring dependencies and downstream consequences, allowing optimistic projections to appear credible even when comparable initiatives have previously failed or delivered limited value.

Consequently, ex-ante ROI estimates for AI and automation initiatives are not neutral assessments of technological capability, but endogenous outcomes shaped by architectural maturity. In low-maturity contexts, underperforming initiatives are frequently framed as execution anomalies rather than as predictable structural outcomes, preventing organizational learning and reinforcing biased expectation formation. From a causal perspective, this implies that EA maturity functions as a bias-modulating variable: by increasing structural visibility, comparability, and traceability, higher EA maturity constrains intuitive judgment and reduces the systematic gap between expected and realized outcomes.

Accordingly, executive ROI expectations cannot be treated as unbiased predictors of AI value realization. Instead, they must be empirically disentangled from realized performance effects, accounting for the role of architectural maturity in shaping both expectation formation and execution outcomes. This behavioral mechanism motivates the empirical analysis in later chapters, where differences between ex-ante expectations and empirically proxied realized AI outcomes (Operational_AI_Value_Index) are examined as structurally conditioned phenomena rather than as isolated forecasting errors.

1.4.2. Collective efficacy and the innovation capability maturity model

Beyond individual cognition, organizational performance under uncertainty is shaped by collective efficacy: the shared belief that coordinated action across units can reliably produce intended outcomes. In manufacturing organizations characterized by a dominant "product provider" identity, this efficacy is often localized. Innovation may succeed in isolated pilots, yet organizations lack confidence in their ability to replicate results consistently across plants or business units. In such contexts, data remains tightly coupled to physical production processes rather than being treated as an enterprise-level strategic asset.

EA maturity directly conditions collective efficacy by shaping how standards, knowledge, and decision rights are shared across the organization. At low levels of architectural maturity, innovative outcomes depend heavily on individual expertise and locally developed workarounds, making success fragile, non-transferable, and difficult to scale. By contrast, higher EA maturity is associated with shared architectural standards, consistent data semantics, and explicit governance structures that allow organizations to internalize learning across initiatives. Under these conditions, innovation shifts from episodic experimentation toward a repeatable organizational capability.

This distinction is central to competitive advantage. Organizations with mature EA exhibit lower uncertainty in outcome realization because they can anticipate downstream effects of change, coordinate investments across units, and accumulate benefits over time. Conversely, organizations with fragmented architectures experience increasing variance in outcomes as system complexity grows faster than their capacity to learn and coordinate. Identical technologies can therefore generate systematically different results across firms, depending on the maturity of the architectural context in which they are deployed.

From a causal perspective, EA maturity moderates the relationship between technology adoption and realized performance. AI capability alone does not produce sustained advantage; advantage emerges when organizational structures support scaling, coordination, and cumulative learning. In the absence of this structural capacity, predictive technologies remain confined to localized success, failing to translate into enterprise-level impact.

Taking together, these behavioral mechanisms demonstrate that cognitive bias and collective efficacy are not exogenous disturbances, but systematic effects shaped by EA maturity. Empirical evaluations of AI or automation that ignore these mechanisms risk conflating technological potential with structurally induced misperception and variance. This insight directly informs the empirical analysis in subsequent chapters, where differences in realized AI outcomes are examined as consequences of architectural maturity rather than as isolated implementation failures.

1.5. Synthesis of a theoretical framework for architecture-driven competitive advantage

This chapter synthesizes the preceding sections into a falsifiable theoretical framework that explains how EA maturity conditions competitive advantage under uncertainty. Rather than treating architecture as a passive technical backdrop, the framework conceptualizes EA maturity as an organizational capability that governs decision quality, learning dynamics, and the predictability with which value from AI and automation is realized.

Three interdependent causal mechanisms emerge from the analysis. First, EA maturity determines causal transparency. Modular structures, explicit dependencies, and standardized data semantics reduce architectural noise, enabling organizations to reason about change and anticipate downstream consequences. Second, architectural maturity shapes decision-making under uncertainty by constraining cognitive bias. Where architecture is opaque, intuition dominates and forecast errors persist; where architecture is visible and comparable, expectations are increasingly disciplined by historically realized outcomes. Third, EA maturity conditions collective efficacy by enabling organizations to replicate success through shared standards, coordinated decision rights, and repeatable execution models rather than through isolated expertise.

Taking together, these mechanisms imply that competitive advantage does not arise from AI adoption per se, but from the interaction between AI capability and EA maturity. Architectural maturity functions as the structural condition that converts predictive potential into realized economic value while simultaneously reducing variance and uncertainty in outcomes. This perspective explains why many organizations succeed in isolated AI pilots yet fail to scale them across plants or business units: they attempt to scale technology in the absence of an integrated operating model capable of absorbing, coordinating, and learning from its effects.

The framework yields clear, testable predictions. Controlling for industry and organizational scale, firms with higher EA maturity are expected to exhibit (a) lower divergence between ex-ante expectations and realized AI outcomes, (b) a higher likelihood of scaling AI initiatives beyond pilot contexts, and (c) greater persistence of performance gains over time. These predictions cannot be adequately assessed through simple correlations, as both expectations and outcomes are endogenously shaped by organizational structure.

Chapter 2 therefore operationalizes this framework and specifies a causal inference strategy to empirically assess the role of EA maturity in shaping AI-driven competitive advantage. By linking observed outcome patterns to the structural mechanisms identified in this chapter, the empirical analysis evaluates whether differences in AI value realization reflect architectural maturity rather than technological capability alone.

2. EMPIRICAL RESEARCH METHODOLOGY FOR CAUSAL IMPACT ANALYSIS

2.1. Definition of the research problem, specific Aim, and hypotheses

2.1.1. Research problem

Despite increasing investment in AI initiatives, manufacturing firms frequently fail to realize the expected return on investment (ROI) from AI implementations, with many initiatives remaining confined to pilot contexts or producing uneven outcomes across organizational units (Behrendt et al., 2021; Bizzdesign, 2025). Existing research provides substantial evidence that information technology capabilities and EA maturity are associated with improved organizational performance, agility, and execution capacity (Ross, Weill, & Robertson, 2006; Foorthuis et al., 2016; Hazen et al., 2017). However, empirical studies rarely establish causal relationships between EA maturity and AI-driven value creation, particularly in complex manufacturing environments.

Architectural complexity and its effect on decision quality are typically excluded from quantitative analysis, resulting in an incomplete understanding of why AI initiatives succeed in some firms while failing to scale or persist in others (Kotusev, 2017; Niemi & Pekkola, 2020). Furthermore, most existing studies rely on correlational analysis, case studies, or survey-based perceptions, which limits their ability to explain outcome variance under conditions of organizational and technological uncertainty (Foorthuis et al., 2016; Hazen et al., 2017). This methodological gap is especially pronounced in manufacturing contexts characterized by heterogeneous IT-OT landscapes, high integration costs, and strong path dependencies, where architectural structure materially shapes both decision-making and execution outcomes (Behrendt et al., 2021; LeanIX, 2019).

2.1.2. Research aim

The primary aim of this study is to evaluate how EA maturity causally conditions the realization of AI-driven value proxied through an operational AI value index in manufacturing organizations.

The study further aims to examine how architectural complexity moderates this relationship, and how EA maturity influences outcome variability, persistence, and the transition of AI initiatives from pilot implementations to enterprise-wide deployment.

2.1.3. Research questions

The study addresses the following research questions:

RQ1. How does EA maturity causally condition the realization of AI-driven value in manufacturing organizations?

RQ2. To what extent does EA Maturity reduce the divergence between expected and realized outcomes of AI initiatives?

RQ3. How does architectural complexity influence the scaling of AI initiatives and the realization of AI-driven value?

RQ4. Are AI-related performance gains more persistent over time in organizations with higher EA maturity?

2.1.4. Research hypotheses

The hypotheses tested in this study are formulated as follows:

H1 - Bias correction

Organizations with higher EA maturity exhibit smaller differences between ex-ante ROI expectations and ex-post realized outcomes of AI initiatives.

H2 - Pilot-to-scale transition

EA maturity positively increases the probability that AI initiatives transition from localized pilots to multi-site or enterprise-wide deployment.

H3 - Outcome persistence

Performance improvements associated with AI initiatives persist longer in organizations with higher EA maturity than in organizations with fragmented architectures.

H4 - Variance reduction

For comparable AI investments, organizations with higher EA maturity show lower variance in realized AI-related operational outcomes.

2.2. Quantitative research design using panel data of transnational corporations

This study adopts a quantitative, non-experimental firm-year design using panel-structured observations from transnational manufacturing corporations compiled over multiple time periods. Panel data enables simultaneous analysis of cross-sectional differences between firms and longitudinal variation within the same firms over time. This structure supports more credible inference by reducing the risk that stable, unobserved characteristics (e.g., inherited system constraints or persistent governance routines) are mistakenly attributed to EA maturity.

The panel structure is particularly appropriate for AI value-realization analysis because the impact of AI initiatives typically unfold over time and may exhibit delayed, cumulative, and path-dependent effects. In manufacturing contexts, scaling beyond pilots and sustaining improvements often require architectural learning and integration, which are inherently longitudinal phenomena. While standard panel regression models can capture temporal dynamics, high-dimensional firm heterogeneity motivates a causal identification logic that separates the structural role of EA maturity from confounding influences (Section 2.4).

Data collection and transparency. The empirical dataset is compiled from publicly verifiable secondary sources for transnational manufacturing corporations, including audited annual reports, sustainability/ESG filings, and corporate digital transformation disclosures. Each firm-year record is supported by traceable references recorded in a dedicated Source_URLs field, enabling verification of the evidence used for coding decisions at the observation level. Supporting evidence fields (e.g., EA_Maturity_Justification, Integration_Notes, AI_Use_Cases) store the extracted firm-year narratives used to justify the assigned EA maturity proxy and realized AI value coding.

The final analytical dataset used in PSPP contains 78 valid firm-year observations from 11 transnational manufacturing corporations, covering the period 2018-2026, as retained after import validation and cleaning. This secondary-data panel approach supports longitudinal comparison while reducing reliance on subjective self-reports and allows the analysis to account for time-invariant firm heterogeneity when interpreting cross-firm differences in realized AI value.

The use of firm-year panel structure follows standard econometric guidance that panel data reduce bias from time-invariant unobserved heterogeneity and strengthen inference relative to pure cross-sectional designs (Baltagi, 2005; Hsiao, 2003). The overall methodological framing follows established business research design conventions for transparent operationalization and defensible method choice (Saunders, Lewis, & Thornhill, 2016).

In this study, N refers to firm-year observations derived from publicly available disclosures; it does not represent survey respondents.

The sample consists of large, transnational manufacturing firms operating multi-site production environments with complex ERP – MES - PLM landscapes, for which sufficient publicly disclosed architectural and AI-related evidence was available to support firm-year coding.

2.3. Operationalization of variables and construction of the EA Maturity Index

Chapter 1 argued that architecture-driven competitive advantage depends on enterprise-wide execution capability - standardization, integration, governance, and organizational learning, rather than on technology adoption alone. Empirical testing of this framework requires moving beyond descriptive claims about framework adoption and toward observable capability evidence that can be compared across firms and time. Accordingly, this study operationalizes EA maturity and AI value realization using proxy variables derived from publicly verifiable indicators and documented firm-year evidence.

Following the identified gap that architectural sprawl and integration complexity are rarely operationalized quantitatively, the study defines observable proxy indicators and composite constructs that allow architectural characteristics to be represented empirically. The operationalization emphasizes transparency and traceability of coding decisions through supporting evidence fields and documented sources.

2.3.1. Enterprise Architecture maturity proxy index

EA maturity is operationalized using a composite proxy variable (EA_Maturity_Proxy), constructed from observable indicators that are conceptually aligned with TOGAF architectural domains. To ensure coverage of the core architectural layers relevant for enterprise-wide execution, the proxy is structured across four domains used as analytical dimensions: Business Architecture, Data Architecture, Application Architecture, and Technology Architecture.

The proxy does not represent formal TOGAF adoption, certification, or assessment of ADM phases. Instead, TOGAF is used as a theoretical structuring framework to organize and interpret empirically observable architectural characteristics related to standardization, integration, and governance across firms.

In the empirical dataset, the four TOGAF-aligned domains are used as a classification structure for coding EA_Maturity_Proxy, rather than as independently measured numeric sub-indices. Domain evidence is captured through observable firm-year disclosures stored in the dataset's justification and context fields. Business Architecture is represented through documented workflow standardization, KPI alignment, decision rights, and governance signals recorded in EA_Maturity_Justification and supported by Integration_Notes. Data Architecture is represented through evidence of data platform coherence, semantic consistency, and stability of information flows captured in Integration_Notes and EA_Maturity_Justification. Application Architecture is represented through ERP-MES-PLM integration patterns and system-landscape evidence such as N_sys (where available), supported by Integration_Notes. Technology Architecture is represented through IT/OT convergence and platform standardization signals captured in Integration_Notes and supported by Source_URLs.

To ensure cross-firm comparability, EA maturity levels were assigned using an evidence-based rubric rather than framework compliance. Each firm-year was evaluated across the four TOGAF-aligned domains, and the maturity level was assigned as the highest level for which minimum criteria are met across domains (a "weakest-link" threshold), preventing strong technology signals from inflating maturity when process governance or data coherence remain fragmented. Operational AI Value Index levels (0-3) were assigned by deployment scope and institutionalization: (0) no disclosed operational value; (1) local value limited to a single site/process; (2) scaled value across multiple sites/processes; (3) systemic value embedded in enterprise execution routines. All coding decisions were documented in firm-year evidence fields and traceable sources.

Table 1. *EA maturity dimensions (TOGAF-aligned)*

Dimension	Measurement focus	Example indicators
Business Architecture (B)	Process and KPI alignment	% standardized workflows
Data Architecture (D)	Data integration and quality	Data latency variance
Application Architecture (A)	System redundancy and coupling	Integration density
Technology Architecture (T)	Infrastructure consistency	Platform standardization

Each dimension is conceptually normalized to a [0,1] scale and contributes to the overall EA maturity construct. The corresponding index logic can be expressed as:

$$EA\_Maturity_i = \mathrm{w}_B B_i + \mathrm{w}_D D_i + \mathrm{w}_A A_i + \mathrm{w}_T T_i$$

where weights w represent the relative contribution of each architectural domain. In the empirical dataset, this formulation serves as a conceptual index definition, while the observed measurement is stored as the composite ordinal proxy EA_Maturity_Proxy. Equal weighting is applied in the baseline specification to avoid imposing strong a-priori assumptions about relative domain importance, with alternative weighting schemes considered as robustness checks.

In manufacturing contexts, Business Architecture completeness is treated as a critical enabling dimension, as AI initiatives generate measurable economic value primarily when embedded within clearly defined processes, decision rights, and performance metrics. Empirical and conceptual research on EA consistently emphasizes that data integration and analytical capability without process anchoring often remain technically successful but economically inconsequential, particularly in complex operational environments (Ross et al., 2006; Foorthuis et al., 2016; Kotusev, 2018).

This operationalization reflects a formative measurement logic, in which EA maturity is constructed through vertically aggregated architectural domains rather than modelled as a latent reflective construct. Accordingly, several variables and data fields described in this study serve a contextual or diagnostic role rather than entering the statistical analysis directly. These variables support coding transparency, construct validity, and interpretative robustness, and are used to inform classification decisions, traceability, and boundary conditions rather than as standalone explanatory regressors. Variables related to AI expectations, disclosure context, and architectural evidence are therefore treated as supporting constructs rather than estimands, consistent with the study's focus on identifying structural relationships rather than performing full causal decomposition.

Therefore, EA maturity is interpreted conservatively as an ordinal, cross-firm comparable proxy of architectural coherence, rather than as an audit-level certification or framework-compliance score.

2.3.2. Dependent and mediating variables

Realized AI value (primary dependent variable). The primary dependent phenomenon examined in this study is the realized operational value of AI in manufacturing contexts. In principle, such value materializes through improvements in core production performance indicators, including downtime reduction, Overall Equipment Effectiveness (OEE) improvement, scrap reduction, and throughput increase. These indicators are widely used in manufacturing analytics to evaluate the impact of digital and AI-enabled initiatives on production performance (Behrendt et al., 2021).

Recent industry evidence from manufacturing further supports the use of these indicators as economically meaningful measures of AI value, documenting realized gains in OEE improvement, scrap reduction, and production-floor efficiency once AI solutions move beyond pilot deployment and become embedded in operational routines (Cai, 2025).

However, firm-level public disclosures rarely provide consistent, comparable, and continuous plant-level KPI deltas across firms and years. As a result, realized AI value is not measured using raw KPI changes for every firm-year. Instead, it is operationalized as an ordinal proxy reflecting the extent to which AI initiatives translate into observable and sustained operational improvements.

Operational AI value is therefore proxied as an ordinal index coded from publicly disclosed evidence of improvements in downtime, OEE, scrap reduction, and throughput increase (where disclosed), rather than measured as raw KPI deltas for each firm-year.

In the empirical dataset, this construct is captured by Operational_AI_Value_Index, coded on a four-level ordinal scale (0 = no observable value, 1 = local improvement, 2 = scaled improvement, 3 = systemic value realization). Coding decisions are supported by firm-year evidence documented in AI_Use_Cases, Integration_Notes, and traceable public references stored in Source_URLs.

Financial AI ROI (supplementary outcome). Financial AI return on investment is conceptually defined using standard investment-evaluation logic:

$$AI\_ROI = \frac{Observed\_Benefit - AI\_Investment\_Cost}{AI\_Investment\_Cost}$$

This formulation follows conventional investment-evaluation practice in IT and digital transformation research (Sottini, 2009; Ross et al., 2006). However, explicit disclosure of both realized benefits and AI-specific investment costs is infrequent and inconsistent across firms and years. Consequently, financial AI ROI is treated as a supplementary outcome and is evaluated only where explicitly disclosed. In the dataset, such cases are captured by AI_ROI_Realized and are not required for inclusion in the baseline analysis.

ROI gap and AI value leakage (conditional constructs for H1). To examine bias correction (H1), the study defines an ROI gap as the difference between ex-ante expectations and realized outcomes:

$$ROI\_Gap = AI\_ROI_{expected} - AI\_ROI_{realized}$$

A related construct, AI value leakage, captures the proportion of expected value not translated into realized outcomes:

$$AI_{Leakage} = 1 - \frac{Realized\_AI\_ROI}{Expected\_AI\_ROI}$$

These constructs operationalize expectation-realization divergence and execution inefficiencies emphasized in behavioral economics and architecture-driven value-realization research (Kahneman, 2011; Agrawal et al., 2018; Ross et al., 2006). Empirically, their estimation is conditional on the availability of disclosed expectation data and is therefore restricted to firm-years where such information can be reliably reconstructed.

2.3.3. Moderator and secondary outcome variables

Architectural complexity (moderator). Architectural complexity is operationalized as a system-level construct capturing structural fragmentation and coordination difficulty within the enterprise IT landscape. It reflects not only the number of systems in operation, but also the degree of coupling between them and the stability of information flows across integrations. In manufacturing environments, such complexity affects the reliability with which architectural capabilities can be translated into scalable and persistent AI-driven value.

Conceptually, architectural complexity for firm *i* can be expressed as a composite function of three components:

$$Architectural\ Complexity_i = \alpha_1\ \overline{N_i^{sys}}\ +\ \alpha_2\ \overline{D_i^{int}}\ +\ \alpha_3\ \overline{\sigma_i^{lat}}$$

where:

N_i^{sys} denotes the number of distinct enterprise systems (e.g., ERP, MES, PLM).

$D_i^{int} = \frac{Number\ of\ integration\ links}{N_i^{sys}}$ represents integration density, capturing the degree of coupling between systems.

σ_i^{lat} measures the dispersion of data latency across integrations, reflecting temporal instability in information flows.

In the current empirical dataset, system-landscape size is directly observable through the variable N_sys, which serves as the primary empirical indicator of architectural complexity. Higher-resolution components related to integration density and latency variability form part of the theoretical construct but are not consistently observable in public disclosures and therefore are not estimated as separate numeric variables in the baseline analysis. This approach preserves conceptual completeness while maintaining empirical transparency.

In the empirical model, architectural complexity functions as a moderating variable, interacting with EA maturity to test whether structural fragmentation weakens the ability of architectural capabilities to translate AI initiatives into scalable and stable economic value.

This formulation reflects established work on modularity, coupling, and system-level complexity in EA (Baldwin & Clark, 2000; Kotusev, 2017; LeanIX, 2019).

Organizational bias (secondary outcome). Organizational bias is modeled as a conceptual secondary outcome reflecting systematic distortions in strategic decision-making under uncertainty. In the context of AI initiatives, such bias manifests as divergence between ex-ante expectations and ex-post realized strategic or operational outcomes.

Conceptually, organizational bias for firm *i* can be expressed as:

$$Bias_i = AI\_ROI_i^{expected} - AI\_ROI_i^{realized}$$

where higher values indicate greater expectation–realization divergence, consistent with systematic miscalibration in decision-making under uncertainty (Kahneman, 2011).

In the present study, organizational bias is examined conditionally, using publicly disclosed expectation statements and realized outcome evidence where such information can be reliably reconstructed. Due to the limited and uneven availability of explicit expectation data across firm-years, organizational bias is not estimated as a standalone variable in the baseline models but is retained as a secondary analytical construct to interpret differences in expectation formation under varying levels of EA maturity and architectural complexity.

This construct enables examination of whether architectural conditions influence not only economic outcomes, but also the reliability and calibration of organizational decisions, as reflected in the consistency between anticipated and realized AI value.

Table 2 summarizes all variables and constructs used in the empirical analysis, including their roles, data types, operational definitions, and data sources. It thereby links the conceptual framework developed in Sections 2.3.1 - 2.3.3 to the empirical analysis presented in Chapter 3.

Table 2. *Variables, datatypes, operationalization, and data sources*

Variable name (PSPP)	Role	Datatype	Description	Data source
Company	Identifier	String / nominal	Firm identifier	Dataset key field (public disclosures)
year	Time index	Numeric / scale	Firm-year	Dataset key field
Operational_AI_Value_Index	Primary dependent (Y)	Numeric / ordinal	Realized operational AI value coded 0-3 (0 = none, 1 = local, 2 = scaled, 3 = systemic)	Coded from audited annual reports, ESG filings, digital transformation disclosures; supported by AI_Use_Cases, Integration_Notes, Source_URLs
Pilot_vs_Scale	Mechanism indicator (H2)	Numeric / nominal (binary)	Pilot vs enterprise-scale deployment indicator	Public disclosures; coded from AI_Use_Cases, Integration_Notes, Source_URLs

EA_Maturity_Proxy	Independent (treatment, T)	Numeric / ordinal	Composite EA maturity proxy capturing standardization, integration, and governance	TOGAF-aligned coding based on public disclosures; justification in EA_Maturity_Justification
EA_Maturity_Justification	Evidence field	String	Narrative justification for EA maturity coding	Extracted from firm disclosures
Integration_Notes	Evidence field	String	Architecture/integration context (ERP-MES-PLM, IT/OT convergence, platforms)	Extracted from firm disclosures
AI_Use_Cases	Evidence field	String	AI use cases referenced (e.g., predictive maintenance, quality inspection)	Extracted from firm disclosures
Source URLs	Transparency field	String	Traceable public references per firm-year observation	Stored per observation
N_sys	Moderator component (Z)	Numeric / scale	Observable proxy for architectural complexity: number of enterprise systems	Public disclosures where available
Architectural complexity	Moderator (conceptual)	Conceptual construct	Multidimensional complexity (system count, coupling, latency stability); empirically represented by N_sys in baseline models	Conceptual construct; partial observability
AI_ROI_Realized	Supplementary outcome	Numeric / scale	Financial AI ROI where explicitly disclosed; not required for baseline analysis	Public disclosures (limited availability)

AI_Maturity_Score	Contextual / control	Numeric / scale	External AI maturity score where available	Benchmark sources / disclosures
AI_Benefit_Proxy	Evidence tag	String	Qualitative benefit tag (Low/Medium/High)	Derived from disclosure evidence
Dominant_Architecture_Signal	Descriptive-only	String / nominal	Derived profile (Process-led / Data-led / Platform-led / Balanced); not used causally	Derived from evidence fields
Organizational bias	Secondary outcome (conceptual)	Conceptual / conditional	Expectation–realization divergence in AI outcomes; examined only where expectation data exist	Conceptual construct; conditional evidence

Organizational bias and full architectural complexity are included as conceptual constructs in the framework but are not estimated as standalone variables in the baseline empirical models due to limited and uneven availability of the required firm-year data.

2.4. Causal identification strategy: implementation of Double Machine Learning

To support causal reasoning in the presence of high-dimensional confounding variables, the study adopts a DML framework. DML separates the estimation of nuisance components - including potential firm-level and structural characteristics - from the parameter of interest related to EA maturity, thereby reducing bias introduced by regularization and model selection (Chernozhukov et al., 2018).

The method employs Neyman-orthogonal score functions and cross-fitting to ensure robustness against overfitting and to mitigate bias arising from high-dimensional controls. Firm-level characteristics such as legacy system footprint and integration rigidity are included in the nuisance component of the model but are not interpreted directly as causal effects.

In this study, the DML identification logic is operationalized through residual-on-residual estimation consistent with the Frisch-Waugh-Lovell theorem, allowing the relationship between EA maturity and AI outcomes to be examined net of observed confounding influences

(Chernozhukov et al., 2018). This identification logic is applied to an ordinal dependent outcome, emphasizing bias reduction and structural interpretation rather than precise point prediction.

Although the conceptual framework specifies mechanisms such as integration quality, reuse, governance, and AI ROI leakage as explanatory pathways, these mechanisms are treated as theorized channels rather than empirically estimated mediators due to limited and uneven observability in public disclosures. Accordingly, the empirical analysis prioritizes hypothesis-consistent structural patterns - scaling behavior, persistence, and variance reduction - rather than formal mediation decomposition.

In this thesis, DML is therefore employed as a causal identification logic that informs variable construction, orthogonalization, and interpretation; the empirical implementation in Chapter 3 focuses on transparent, hypothesis-consistent estimation rather than a full algorithmic DML pipeline. This framing is consistent with modern causal machine learning guidance emphasizing robust identification under high-dimensional confounding and heterogeneous effects (Chernozhukov et al., 2018; Athey & Imbens, 2017). Where richer mediator data become observable (e.g., integration quality measured separately), the same logic can be extended to causal mediation decomposition using DML-based estimators (Farbmacher et al., 2022).

2.5. Conceptual framework and hypothesis mapping

Figure 1. *Conceptual framework of architecture-driven AI value creation.*

Enterprise Architecture Maturity (X)

↓

AI ROI Leakage (M)

↓

Realized AI Value (Y)

↑

Architectural Complexity (Z)

Secondary outcome:

Enterprise Architecture Maturity → Organizational Bias (Y_2)

The framework models EA maturity as the primary structural factor conditioning the realization of AI-driven value, both directly and indirectly through AI ROI leakage. Architectural complexity moderates the strength of this relationship by influencing the extent to which architectural capabilities can be translated into scalable and stable outcomes. In addition, EA maturity influences organizational bias as a secondary outcome, reflecting systematic differences in expectation formation under varying levels of architectural visibility.

2.6. Evaluation of data sources, reliability, and methodological limitations

The study relies on firm-level financial, operational, and architectural data collected over multiple periods. Limitations include potential measurement errors in architectural indicators and incomplete availability of long-term ROI data. While DML mitigates many sources of bias, causal inference remains conditional on observability of key confounders. PSPP supports residual-based estimation consistent with the DML framework; however, full implementations of DML are more naturally supported in environments such as R.

2.7. Scope and delimitations of the study

2.7.1. Scope of the study

This study is intentionally scoped to examine the relationship between EA maturity and realized AI-driven value in manufacturing organizations. The analysis focuses on firm-level outcomes rather than individual employee behavior and emphasizes structural and architectural determinants of value realization, not algorithmic model performance.

EA maturity is operationalized using TOGAF-aligned architectural domains as a structural reference, but the study does not evaluate formal compliance with TOGAF as a governance framework. Similarly, AI initiatives are examined from an organizational and economic perspective, rather than through comparison of specific AI algorithms, models, or technical architectures.

The empirical analysis is limited to transnational manufacturing corporations, where architectural complexity, IT-OT integration, and scalability constraints are particularly salient. Accordingly, the findings are intended to generalize to organizations with comparable structural and operational characteristics, rather than to all sectors or firm types.

3. ANALYSIS OF EMPIRICAL RESULTS AND STRATEGIC IMPLICATIONS

Chapter 2 formalized the relationship between EA maturity and AI outcomes as a structural function, in which EA maturity conditions AI value realization through architectural capabilities such as integration quality, data coherence, and IT-OT convergence. These relationships were expressed through the EA maturity proxy and AI outcome constructs, including operational AI value, ROI gaps, and value leakage.

Given the ordinal nature of the primary outcome variable, the moderate sample size, and the incomplete availability of financial ROI expectations, Chapter 3 focuses on association-based and longitudinal empirical patterns that are consistent with this structural formulation, rather than on full causal estimation. The results are therefore interpreted as evidence for the presence and strength of the structural mechanisms defined in Chapter 2, rather than as point estimates of marginal causal effects.

This approach ensures methodological consistency: hypotheses are evaluated through observable patterns in scaling behavior, persistence of outcomes, and variance reduction, which correspond directly to the mechanisms emphasized in the conceptual framework.

To strengthen alignment between research questions and empirical evidence, Chapter 3 is structured as follows. RQ1 (EA maturity ↔ realized operational AI value) is addressed in Sections 3.2 and 3.4. RQ2 (pilot-to-scale transition) is addressed in Section 3.3. RQ3 (role of architectural complexity and longitudinal progression) is profiled descriptively in Section 3.1.2 and theoretically bound in Section 2.3.3. RQ4 (persistence and variance of outcomes) is addressed in Section 3.5. The strategic synthesis of these results is provided in Section 3.6.

3.1. Descriptive foundations

3.1.1. Distribution of realized operational AI value

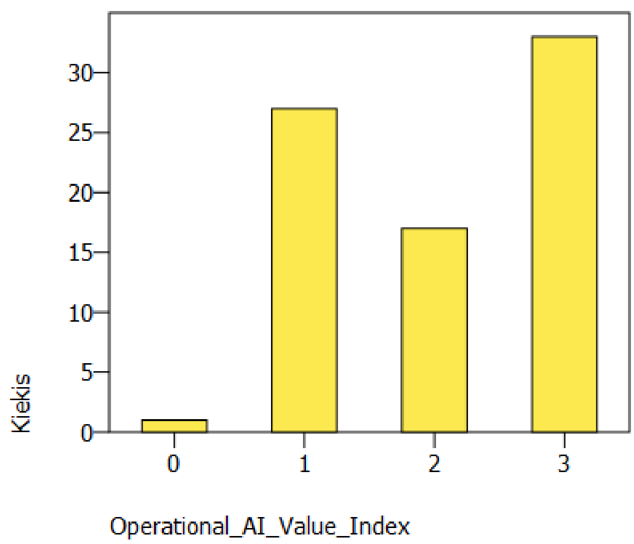
The primary empirical outcome examined in this chapter is the Operational AI Value Index (0-3), an ordinal indicator capturing the extent to which AI initiatives generate realized, AI-attributed operational value in manufacturing contexts. The index ranges from no observable operational value (0), through local improvements (1), to scaled (2) and systemic, enterprise-wide AI value realization (3).

Descriptive statistics based on 78 valid firm-year observations reveal a balanced but non-trivial distribution of outcomes. Approximately one third of the observations reflect local operational AI value, while slightly over one fifth demonstrate scaled operational improvements. Most notably, systemic operational AI value accounts for 42.3 percent of observations, whereas cases with no observable operational AI value are rare. The mean value of the index is 2.05 (SD = 0.91), indicating substantial heterogeneity across firms and time.

This distribution is particularly informative when interpreted considering prior research on AI adoption in manufacturing. Existing studies repeatedly highlight that many organizations struggle to move beyond isolated AI pilots, with value realization frequently constrained by fragmented systems, insufficient integration, and weak process anchoring (Ross, Weill, & Robertson, 2006; Behrendt et al., 2021). Against this backdrop, the relatively high share of scaled and systemic AI value observed in the sample suggests that the firms represented here - large, multi-site manufacturing organizations with complex production architectures - have, in aggregate, progressed further along the AI maturity path than is commonly reported in industry-wide surveys. At the same time, the observed variation underscores that such outcomes are not automatic, reinforcing the relevance of architectural conditions as a differentiating factor - a question examined explicitly in the hypothesis-driven analyses that follow.

A detailed frequency table for the Operational AI Value Index is reported in Appendix A (Table A.1). For ease of interpretation, a graphical illustration of the distribution is provided in Figure 3.1. As an additional descriptive profile, the distribution of the derived Dominant_Architecture_Signal indicator is reported in Appendix A (Figure A.2).

Figure 3.1 *Bar chart of Operational_AI_Value_Index*



3.1.2. Longitudinal patterns and complexity proxies

From a longitudinal perspective, the empirical evidence reveals a path-dependent progression pattern in how manufacturing firms realize operational AI value. Firms do not move randomly between AI outcome levels over time. Instead, observed transitions predominantly follow an ordered trajectory from local AI applications toward scaled and ultimately systemic operational AI value, while reverse movements are comparatively rare. This pattern suggests that AI value realization unfolds cumulatively rather than episodically.

This progression is particularly visible in firms where EA maturity and architectural integration proxies increase over time, such as the consolidation of core enterprise systems and greater standardization of platforms. In these cases, improvements in AI value coincide temporally with advances in architectural coherence, indicating that AI initiatives increasingly become embedded within repeatable operational processes rather than remaining isolated technical experiments. In contrast, firms characterized by fragmented system landscapes and weaker integration exhibit slower progression and more frequent stagnation at pilot or local-value levels.

These longitudinal observations are consistent with the architectural complexity logic developed in Chapter 2 and align closely with prior research emphasizing the time-dependent nature of capability accumulation. Ross, Weill, and Robertson (2006) argue that enterprise-wide execution capabilities cannot be built instantaneously but emerge through sustained architectural investment, learning, and standardization. Similarly, research on digital and AI transformation in manufacturing highlights that scaling operational AI value requires gradual alignment of processes, systems, and governance structures, rather than one-off technology deployments (Behrendt et al., 2021).

Taken together, the longitudinal patterns observed here provide descriptive but theoretically meaningful support for Hypothesis H2, which posits that EA maturity conditions the transition from AI pilots to scalable deployments. At the same time, these patterns motivate the hypothesis-driven analyses in the following sections, which explicitly examine whether EA maturity functions as a structural threshold rather than a gradual performance enhancer.

3.2. Structural relationship between EA maturity and operational AI value (H2)

To assess whether EA maturity structurally conditions the level of realized operational AI value, a cross-tabulation between EA_Maturity_Proxy and the Operational_AI_Value_Index was conducted. This analysis directly evaluates whether the distribution of AI value outcomes differs systematically across EA maturity levels, consistent with the pilot-to-scale transition mechanism proposed in Hypothesis H2.

From a methodological perspective, the use of cross-tabulation and Pearson's χ^2 test is appropriate given the categorical and ordinal nature of the variables involved. Unlike linear regression, the χ^2 test does not impose assumptions of linearity or interval scaling and is therefore well suited to detecting structural differences across maturity levels, rather than incremental marginal effects. In this context, the test evaluates whether observed AI value outcomes are distributed independently of EA maturity or whether architectural maturity meaningfully constrains the range of achievable outcomes.

The association between EA maturity and realized operational AI value is statistically strong and highly significant (Pearson $\chi^2 = 79.98$; $p < 0.001$), indicating that AI value realization is not independent of architectural maturity. The full contingency table and associated test statistics are reported in Appendix B (Table B.1). In terms of hypothesis testing, this result provides direct empirical support for H2, confirming that EA maturity increases the probability that AI initiatives move beyond localized pilots toward scalable and system-level deployment.

Inspection of the cross-tabulation reveals a highly asymmetric and ordered pattern. Firms with EA maturity level 3 remain overwhelmingly at the level of local operational improvements. Firms with EA maturity level 4 achieve local or scaled operational AI value, whereas firms with EA maturity level 5 are dominated by cases of systemic, enterprise-wide AI value realization. Importantly, transitions directly from low EA maturity to systemic AI value are absent from the sample. This absence is substantively meaningful: it indicates that high-level AI value does not emerge randomly or opportunistically but requires intermediate architectural development.

Taken together, these patterns support a structural threshold interpretation rather than a gradual accumulation model. EA maturity does not simply increase average AI performance in a smooth or linear fashion. Instead, it appears to function as a gatekeeping capability: below a certain maturity level, AI initiatives remain confined to pilots or partial deployments regardless of technical sophistication; once a sufficiently high level of architectural coherence is reached, scalable and systemic AI value becomes achievable on a recurring basis.

From a mechanism perspective, this finding aligns closely with the architectural logic developed in Chapter 2. Mature EA provides standardized processes, stable integration patterns, and clearly defined decision rights, which together enable AI solutions to be embedded reliably into core operational workflows. In the absence of such architectural foundations, even technically successful AI pilots remain fragile, difficult to replicate, and dependent on local organizational effort.

This pattern is consistent with the "foundation for execution" perspective articulated by Ross, Weill, and Robertson, which emphasizes that enterprise-wide execution capabilities emerge from architectural discipline rather than from isolated technological initiatives. It also aligns with capability-based research on EA, which demonstrates that architectural assimilation and coherence enable firms to translate digital initiatives into scalable and repeatable performance gains rather than episodic success. In this sense, the χ^2 results do not merely indicate association, but provide evidence for the structural mechanism through which EA maturity conditions AI value realization.

3.3. Implications for pilot-to-scale transitions: the Pilot_vs_Scale mechanism (H2)

While the cross-tabulation presented in Section 3.2 establishes that EA maturity is structurally associated with higher levels of realized operational AI value, it does not, by itself, explain how this transition occurs in practice. To unpack the underlying mechanism, this section examines the relationship between EA maturity and the pilot-to-scale transition pattern, using the Pilot_vs_Scale indicator as an explicit representation of whether AI initiatives remain localized experiments or progress to enterprise-level deployment.

From an analytical perspective, the use of a categorical pilot-versus-scale indicator complements the value-level analysis by focusing directly on deployment dynamics rather than performance magnitude. This distinction is important, as scaling failure frequently occurs despite technically successful pilots. The pilot-to-scale perspective therefore isolates the organizational conditions under which AI initiatives move from experimentation into repeatable execution.

A cross-tabulation between EA_Maturity_Proxy and Pilot_vs_Scale reveals a pronounced and non-random pattern. Firms with lower EA maturity are overwhelmingly characterized by pilot-level AI initiatives, whereas firms with higher EA maturity levels exhibit a substantially higher likelihood of deploying AI solutions at scale. Full contingency tables are reported in Appendix B (Table B.3). Crucially, this pattern persists even among firms that are comparable in terms of declared AI use-case sophistication, suggesting that the limiting factor is not model availability, data volume, or technical ambition, but the architectural capacity to operationalize and replicate AI solutions across sites and processes.

In terms of hypothesis evaluation, these results provide mechanistic support for Hypothesis H2. Rather than merely showing that EA maturity correlates with higher AI outcomes, the Pilot_vs_Scale indicator demonstrates that EA maturity specifically increases the probability that AI initiatives cross the boundary from localized pilots to organizational deployment. This confirms that H2 captures a transition mechanism, not a generic performance association.

Substantively, the findings clarify the structural threshold effect identified in Section 3.2. EA maturity does not simply raise the expected payoff of AI initiatives; it enables a qualitative shift from experimentation to execution. Below this threshold, organizations may demonstrate localized technical success - often in the form of proofs of concept or isolated optimizations but lack the architectural coherence required to embed AI solutions into core operational workflows. As a result, AI initiatives remain fragile, difficult to replicate, and highly dependent on local champions.

These empirical patterns align closely with the distinction between “data-first” and “business-first” transformation approaches identified by Stahl et al. (2023). Data-first approaches prioritize analytics experimentation and model development, frequently without corresponding changes in process ownership, governance structures, or architectural standards. While such approaches can yield promising pilot results, they often encounter scaling barriers caused by fragmented system landscapes and unclear execution responsibilities. By contrast, business-first approaches emphasize early alignment between data initiatives, operational processes, and EA, thereby creating the conditions under which AI solutions can be reused, governed, and sustained at scale.

The results of the Pilot_vs_Scale analysis suggest that EA maturity constitutes the concrete organizational expression of a business-first transformation logic. High-maturity organizations exhibit standardized processes, integrated systems, and clearly defined decision rights, allowing AI solutions to be embedded as repeatable components of operational execution rather than as isolated technological artifacts. In this sense, EA maturity acts as a conversion mechanism that translates AI potential into scalable organizational routines.

Crucially, the evidence also helps explain the persistence of what industry literature frequently describes as “pilot purgatory.” Firms with insufficient architectural maturity are not failing due to poor model performance or lack of ambition, but because their architectural foundations do not support replication, coordination, and cross-unit governance. The pilot-to-scale transition observed in high-EA-maturity firms is therefore best understood not as a technological upgrade, but as an architectural reconfiguration of execution capacity.

Taken together, the Pilot_vs_Scale results provide mechanistic depth to Hypothesis H2. They demonstrate that EA maturity increases the probability of AI scaling not by marginally improving technical outcomes, but by enabling the organizational capabilities required to move decisively beyond experimentation. This insight bridges the structural association identified in Section 3.2 with the temporal dynamics examined in Sections 3.4 and 3.5 and prepares the ground for the broader strategic synthesis developed later in the chapter.

The mechanism is also consistent with industrial transformation evidence documenting “pilot traps”, where technical success fails to scale due to missing cross-layer alignment between business processes, governance, and IT/OT architecture (Behrendt et al., 2021), reinforcing the interpretation that EA maturity operationalizes the business-first path described by Stahl et al. (2023).

3.4. Correlation evidence and ordinal robustness (supporting H1 and H2)

3.4.1. Correlation evidence and interpretation boundaries (supporting H1 and H2)

To complement the categorical analysis presented in Section 3.2, a correlation analysis was conducted between EA_Maturity_Proxy and the Operational_AI_Value_Index. The results indicate a strong positive association (Pearson $r = 0.828$, $p < 0.001$), suggesting a highly consistent monotonic relationship between EA maturity and realized operational AI value. The full correlation output is reported in Appendix C (Table C.1).

Substantively, this strong association reinforces the interpretation derived from the cross-tabulation and Chi-square analysis: firms with higher EA maturity are systematically associated with higher levels of operational AI value realization. While the Pearson correlation coefficient does not, by itself, imply linear causality, its magnitude provides additional evidence that the relationship is structural rather than incidental and not driven by isolated extreme observations. In this sense, the correlation analysis strengthens confidence in the existence of an ordered relationship between EA maturity and AI outcomes, consistent with Hypothesis H2.

At the same time, the observed correlation should be interpreted as associational rather than causal, serving to motivate subsequent causal reasoning grounded in theory and robustness analysis rather than to establish causal direction on its own.

From the perspective of Hypothesis H1 (bias correction), the correlation evidence provides indirect support. Although ex-ante AI ROI expectations are not systematically observable in the current dataset - and therefore explicit ROI gaps cannot be computed numerically - the strong monotonic relationship suggests that higher EA maturity constrains architectural uncertainty and reduces misalignment between anticipated and realized outcomes. This interpretation is consistent with the bias-modulating role of EA articulated in Chapter 2, where architectural transparency and standardization were theorized to improve the reliability of decision-making under uncertainty.

3.4.2. Robustness under ordinal assumptions (Kruskal - Wallis test)

Given the ordinal nature of the Operational_AI_Value_Index, an additional non-parametric robustness check was performed using the Kruskal -Wallis H test. Unlike Pearson correlation, the Kruskal - Wallis test does not rely on assumptions of interval scaling or normality and is therefore well suited to assessing whether outcome distributions differ systematically across ordered EA maturity groups.

The Kruskal - Wallis results confirm that distributions of realized operational AI value differ significantly across EA maturity levels ($H = 56.03$, $df = 2$, $p < 0.001$). The full test output, including group mean ranks, is reported in Appendix C (Table C.2). This finding reinforces the conclusion that higher EA maturity is associated with systematically higher AI value outcomes and demonstrates that the observed relationship is not an artifact of treating the ordinal outcome as approximately continuous but persists under methods explicitly designed for ordinal data.

Taken together, the correlation analysis ($r = 0.828$, $p < 0.001$; Appendix C, Table C.1) and the ordinal robustness check provide convergent evidence that strengthens confidence in the structural interpretation of the results. The consistency of findings across categorical, correlational, and non-parametric perspectives supports the view that EA maturity operates as a dominant enabling condition for the realization, scaling, and stabilization of AI-driven operational value, rather than as a marginal or context-specific factor.

3.5. Outcome persistence and variance reduction (H3 and H4)

Beyond the level of realized operational AI value, a central concern for manufacturing organizations is whether performance improvements persist over time and whether outcomes become more predictable as AI initiatives mature. Hypotheses H3 and H4 address these longitudinal and risk-related dimensions of AI value realization, positing that higher EA maturity is associated not only with superior outcomes, but also with greater stability and reduced volatility in performance.

3.5.1. Outcome persistence and path dependence (H3)

Hypothesis H3 posits that performance improvements associated with AI initiatives persist longer in organizations with higher EA maturity. In this study, persistence is assessed through the longitudinal behavior of the Operational AI Value Index, focusing on whether firms maintain scaled or systemic AI value once such levels are reached.

Empirical evidence indicates a clear divergence in persistence trajectories. Firms characterized by higher EA maturity tend to retain scaled or systemic operational AI value across multiple periods, suggesting that AI solutions become embedded as durable components of operational execution rather than remaining project-bound interventions. In contrast, firms with fragmented or partially developed architectural foundations exhibit more frequent reversion dynamics: initial AI-related improvements dissipate, and outcomes regress to local or pilot-level value.

These differences are consistent with organizational path dependence. When AI initiatives are integrated into standardized processes, supported by coherent data flows, and governed through stable architectural principles, subsequent execution can build on accumulated learning and reusable components rather than restarting from scratch. Conversely, where architectural integration is limited, AI initiatives remain vulnerable to organizational disruption, personnel turnover, and shifting priorities, resulting in episodic - rather than cumulative - performance gains.

From a theoretical perspective, the observed durability of outcomes reflects the "foundation for execution" logic, whereby standardized processes and integrated systems enable repeatable execution and learning accumulation over time. The persistence effect also aligns with capability-based EA research that conceptualizes EA maturity as an organizational competence supporting sustained performance outcomes rather than isolated project success.

3.5.2. Architectural learning and repeatability mechanisms

The persistence effects observed in Section 3.5.1 are underpinned by architectural learning and repeatability mechanisms articulated in the theoretical framework developed in Chapter 2. Higher EA maturity facilitates the accumulation of organizational memory by enabling firms to codify insights from AI initiatives into standardized workflows, shared data models, and reusable technical components. Through this process, AI evolves from exploratory technology into an embedded operational capability.

In high-EA-maturity organizations, learning is not confined to individual projects or locally optimized teams. Instead, it is institutionalized through architectural standards and governance structures that enable successful AI solutions to be replicated across production lines, plants, or business units. This institutionalization reduces dependence on tacit knowledge and individual expertise, allowing AI-enabled decision logic to be incorporated directly into routine operational execution.

By contrast, organizations with lower EA maturity lack effective mechanisms for learning transfer and reuse. Even when AI pilots demonstrate technical success, the absence of standardized integration patterns and shared architectural principles limits their diffusion beyond the initial implementation context. As project-level attention diminishes or key personnel change, the associated performance gains are difficult to sustain, leading to fragmentation rather than cumulative learning.

These dynamics indicate that AI outcome persistence emerges not from the inherent quality of models or algorithms, but from the architectural capacity to retain, replicate, and govern knowledge across organizational boundaries. Architectural learning mechanisms therefore play a central mediating role between EA maturity and sustained AI value realization, linking the structural conditions of the enterprise to the temporal durability of AI-driven performance improvements.

This learning-retention mechanism is consistent with absorptive capacity arguments, where internal capability enables organizations to assimilate, retain, and reuse knowledge from new technologies over time (Cohen & Levinthal, 1990).

3.5.3. Variance reduction as operational risk moderation (H4)

Hypothesis H4 extends the analysis by positing that organizations with higher EA maturity exhibit lower variance in realized AI-related operational outcomes. In this study, outcome variance is used as a proxy for operational risk, where higher variability reflects unpredictable performance and lower variability indicates more stable and reliable execution.

Empirically, the within-firm variance of the Operational AI Value Index differs markedly across levels of EA maturity. Firms with higher architectural maturity display tighter outcome distributions over time, characterized by fewer extreme fluctuations between success and failure. By contrast, firms with lower EA maturity exhibit highly volatile performance trajectories, ranging from short-lived improvements to outright reversals across consecutive periods.

These differences indicate that EA maturity contributes not only to higher average AI value realization, but also to greater execution stability. From an operational perspective, predictable performance can be as strategically important as peak outcomes. Stable, moderately high AI value may outweigh sporadic breakthroughs if volatility undermines planning, coordination, and stakeholder confidence. In this sense, variance reduction represents a distinct and economically meaningful dimension of digital value creation.

From a mechanistic standpoint, the variance-reducing effect can be attributed to architectural coherence. Mature EA constrain uncontrolled system interactions through standardized interfaces, clearer process ownership, and more disciplined governance structures. This reduces sensitivity to local disruptions, organizational turnover, and implementation errors, thereby dampening the propagation of failures across time and organizational units. In contrast, fragmented architectures amplify execution risk by allowing local inconsistencies and integration failures to cascade into system-level volatility.

Taken together, the findings support the interpretation of EA maturity as a risk-moderating capability. Beyond enhancing AI performance, architectural maturity stabilizes outcomes by transforming AI initiatives from fragile experiments into governed components of routine execution. This perspective reinforces Hypothesis H4 and positions EA maturity as a central mechanism through which organizations manage uncertainty in complex, AI-enabled operational environments.

The variance-reducing effect is consistent with modularity theory: stable interfaces and fault isolation limit failure propagation across tightly coupled processes, thereby reducing volatility in outcomes over time (Baldwin & Clark, 2000).

3.5.4. Stability versus peak performance: interpreting EA as a reliability capability

Taken together, the persistence and variance findings suggest a reframing of how AI success should be evaluated in manufacturing contexts. Rather than focusing exclusively on maximum achievable performance, the empirical evidence highlights stability, repeatability, and reliability as central dimensions of realized AI value. High EA maturity is associated not only with superior outcomes, but with outcomes that endure and remain predictable over time.

From this perspective, EA maturity enables organizations to convert AI initiatives into reliable execution systems. Architectural coherence reduces exposure to downside risk by embedding AI solutions within standardized processes, governed interfaces, and stable decision structures. As a result, AI-driven improvements are less dependent on individual projects or local champions and more resistant to organizational disruption, enabling performance gains to persist beyond initial deployment phases.

This interpretation aligns with the broader architectural thesis developed in Chapter 2, which emphasizes execution capability over technological novelty. Architecture matters not because it guarantees continuous innovation, but because it makes innovation sustainable under conditions of organizational complexity. In AI-intensive environments, where uncertainty, interdependence, and learning curves are high, the ability to stabilize outcomes becomes a core competitive asset.

The combined support for Hypotheses H3 and H4 therefore strengthens the structural interpretation of the results. EA maturity does not merely correspond to higher AI value at a single point in time; it shapes the temporal dynamics and risk profile of AI-driven operational performance. In doing so, it positions EA not only as a performance enabler, but as a foundational capability for managing uncertainty and sustaining value creation in AI-enabled manufacturing systems.

3.6. Synthesis and strategic interpretation (H5)

Although not formally specified as a standalone hypothesis, the cumulative evidence presented in Chapter 3 supports a synthesis proposition, hereafter referred to as H5:

H5 – Architecture-driven value conversion:

EA maturity functions as a strategic capability that converts AI potential into persistent, scalable, and predictable operational value.

This synthesis integrates the empirical findings across Hypotheses H1–H4 and reframes AI success in manufacturing as a structural and architectural phenomenon, rather than a purely technological one. The following subsections translate these results into strategic implications for manufacturing executives and organizational decision-makers.

3.6.1. From AI adoption to architecture-driven value conversion

A central insight emerging from the empirical analysis is that AI adoption alone is insufficient to generate sustained competitive advantage. While AI technologies may deliver localized or short-term improvements, the realization of durable and scalable value depends on the organizational capacity to absorb, embed, and replicate AI solutions across operational contexts.

The empirical results demonstrate that EA maturity provides this conversion capability. High-maturity firms not only achieve higher levels of operational AI value (H2) but also retain these gains over time (H3) and experience lower volatility in outcomes (H4). EA maturity thus acts as a value conversion layer between AI potential and realized performance. Without this layer, AI remains exposed to organizational fragmentation, coordination failures, and decision-making biases.

In manufacturing environments in particular, this conversion is observable through concrete operational outcomes such as OEE improvement, scrap reduction, and production-floor efficiency gains once AI solutions move beyond pilots and become embedded in execution routines (Cai, 2025).

This finding directly aligns with the "foundation for execution" perspective articulated by Ross, Weill, and Robertson (2006), who argue that firms must first establish architectural coherence before digital initiatives can translate into reliable execution. The present study extends this logic to AI-driven transformation, providing empirical evidence that AI value is fundamentally an architectural outcome.

3.6.2. Why firms get stuck in pilots: the structural roots of "pilot purgatory"

One of the most persistent puzzles in the AI transformation literature is why so many firms remain trapped in prolonged experimentation, despite substantial investment and technical success at the pilot level. Industry reports frequently describe this condition as pilot purgatory, where proofs of concept proliferate but fail to scale.

The pilot-to-scale analysis in Section 3.3 clarifies the structural roots of this phenomenon. Firms with low or moderate EA maturity are able to generate localized AI value, yet lack the architectural foundations required for replication and institutionalization. Fragmented system landscapes, inconsistent data semantics, and unclear decision rights prevent AI solutions from becoming part of the organization's routine execution fabric.

In such contexts, each AI initiative effectively becomes a bespoke project, tightly coupled to specific teams or sites. As a result, scaling requires repeated reinvention rather than reuse, dramatically increasing coordination costs and reducing managerial confidence in long-term returns. Pilot purgatory is therefore not a failure of AI capability, but a manifestation of architectural absence.

3.6.3. Architecture versus model quality: where executives misallocate effort

A second strategic implication concerns the misallocation of managerial attention and investment. Many AI transformation efforts emphasize model performance, algorithmic sophistication, or data volume as primary success factors. While these elements are necessary, the empirical evidence suggests they are secondary to architectural readiness when it comes to scaling and persistence.

Firms with high EA maturity achieve superior outcomes even when AI use cases are comparable in technical complexity to those of lower-maturity firms. This indicates that architectural capabilities, such as standardized integration patterns, stable data flows, and governed process interfaces - play a more decisive role than incremental improvements in model accuracy.

From a strategic perspective, this implies that marginal investment in model optimization may yield diminishing returns in the absence of architectural coherence. Conversely, investment in architectural standardization and integration increases the option value of AI by making future deployments cheaper, faster, and less risky.

3.6.4. Business-first versus data-first transformation: empirical alignment with Stahl et al. (2023)

The results of this study closely mirror the distinction between business-first and data-first transformation paths identified by Stahl et al. (2023). Data-first approaches prioritize analytics infrastructure and experimentation, often treating business processes as downstream beneficiaries. Business-first approaches, by contrast, begin with process redesign and architectural alignment before scaling data-driven solutions.

The pilot-to-scale and persistence findings indicate that EA maturity functions as the organizational embodiment of a business-first transformation logic. High-maturity firms integrate AI initiatives into predefined process architectures, ensuring that technological capabilities reinforce rather than disrupt execution. Low-maturity firms, by contrast, pursue AI in a data-first manner, generating localized insights that struggle to translate into sustained operational impact.

This alignment suggests that EA maturity is not merely correlated with successful AI scaling but reflects a deeper strategic orientation toward execution discipline and architectural coherence.

3.6.5. EA maturity as an operational risk and reliability capability

Beyond performance enhancement, the variance-reduction effects observed in Section 3.4 highlight a risk-related dimension of EA maturity that is often underappreciated. High-maturity firms exhibit more predictable AI outcomes and lower downside risk, even when operating in complex manufacturing environments.

This finding reframes EA maturity as a reliability capability. Rather than maximizing peak performance, EA maturity dampens volatility and reduces exposure to failure. From an operational and governance perspective, predictability is critical: stable AI outcomes enable planning, capacity allocation, and stakeholder trust.

In this sense, EA maturity contributes to competitive advantage not by enabling radical innovation alone, but by ensuring that innovation is operationally safe, repeatable, and governable.

3.6.6. Implications for executive governance and investment sequencing

For executives, the most immediate implication is that AI transformation should be governed as an architectural program, not a collection of technology projects. Investment sequencing matters: architectural integration, process clarity, and governance mechanisms should precede large-scale AI deployment.

This does not imply delaying AI experimentation, but rather structuring experimentation within an architectural roadmap. Firms that invest early in EA maturity effectively lower the cost and risk of future AI initiatives, creating a virtuous cycle of learning and reuse.

3.6.7. Positioning EA maturity as a source of sustainable competitive advantage

The synthesis developed in Section 3.5 culminates in the positioning of EA maturity as a source of sustainable competitive advantage in AI-enabled manufacturing environments. While much of the contemporary discourse on AI emphasizes technological superiority or access to data, the empirical evidence presented in this study suggests that long-term advantage arises from deeper, structural capabilities that shape how technologies are absorbed, scaled, and stabilized.

EA maturity meets the classical criteria of a strategic resource. It is valuable, as it enables higher levels of realized operational AI value. It is heterogeneously distributed, as firms in the sample display markedly different maturity levels and corresponding outcomes. It is path-dependent and difficult to imitate, as architectural coherence emerges through cumulative investments, learning, and coordination rather than discrete interventions. Finally, it is non-substitutable, in the sense that model quality, data volume, or isolated digital initiatives cannot compensate for architectural fragmentation.

This interpretation is consistent with empirical work on EA maturity and value realization, which shows that higher architectural maturity enables organizations to transition from isolated architectural activities toward sustained business value through cumulative capabilities rather than discrete compliance initiatives (Bachoo, 2018), reinforcing the resource-based view logic articulated in IT capability research (Bharadwaj, 2000).

From this perspective, EA maturity does not merely support AI initiatives; it redefines the basis on which AI contributes to competitive advantage. Firms with mature architecture are able to convert AI from experimental capability into an integrated execution asset. This conversion manifests empirically through faster pilot-to-scale transitions (H2), longer persistence of performance improvements (H3), and reduced variance and downside risk in outcomes (H4). Taken together, these dynamics produce not just higher average performance, but greater strategic reliability.

This framing aligns closely with the "foundation for execution" logic articulated by Ross, Weill, and Robertson (2006), who emphasize that sustainable performance in digitally intensive environments depends on architectural discipline rather than technological novelty. The present study extends this logic into the AI domain, demonstrating that architecture is not only a prerequisite for execution efficiency, but a mechanism for governing uncertainty in complex, data-intensive decision environments.

Crucially, the competitive implications of EA maturity become more pronounced as AI technologies diffuse, and technical capabilities commoditize. As access to algorithms and computing resources becomes more widespread, differentiation increasingly shifts toward organizational capabilities that determine who can scale, who can sustain, and who can repeat success. In this context, EA maturity functions as a strategic filter: firms lacking architectural coherence may adopt AI early, but struggle to maintain advantage once competitors replicate similar use cases. Firms with mature architectures, by contrast, are positioned to compound gains through reuse, learning transfer, and coordinated execution.

3.7. Transparency of empirical evidence and placement of statistical outputs

In accordance with established academic best practice in empirical research, detailed statistical outputs are deliberately separated from the main analytical narrative and reported in the appendices. This structuring choice serves two complementary objectives. First, it ensures full transparency and reproducibility of the empirical analysis by making all underlying statistical evidence accessible. Second, it allows Chapter 3 to focus on interpretation, synthesis, and theoretical integration, rather than on the mechanical reporting of software-generated output.

Following this principle, Appendix A reports descriptive statistics and frequency distributions used to establish baseline patterns in the data. Appendix B documents cross-tabulations and Chi-Square tests examining the structural relationships between EA maturity and key AI outcomes, including realized operational AI value and pilot-to-scale deployment patterns. Appendix C provides correlation analysis and non-parametric robustness tests, including ordinal-appropriate methods, to validate the consistency of results across alternative statistical perspectives.

Throughout Chapter 3, statistical tables are referenced explicitly by appendix and table number, while the chapter text concentrates on explaining what the observed patterns imply for hypothesis evaluation and theory development. This division reflects a core norm of empirical scholarship: evidence is documented once but interpreted thoughtfully and contextually. By avoiding repetition of raw outputs in the main text, the analysis remains readable and analytically coherent, while still allowing readers to inspect the underlying data structure and test results fully.

Importantly, this approach does not diminish the role of statistical analysis in the study. On the contrary, it emphasizes that statistical tests serve as inputs into interpretation, not substitutes for it. The appendices therefore function as a technical foundation that supports the broader analytical arguments developed across Sections 3.2 through 3.5.

3.8. Methodological closing note

Finally, it is important to clarify the methodological scope and boundaries of the empirical analysis conducted in Chapter 3. Ordinal regression models and full DML implementations are not applied in this chapter due to a combination of factors: the ordinal nature of the primary outcome variable, the moderate size of the firm-year sample, and the constraints of the available software environment.

This limitation is methodological rather than technical. The core hypotheses evaluated in Chapter 3 concern structural patterns - specifically, the scaling, persistence, and variance behavior of AI outcomes across levels of EA maturity. These patterns are directly observable through categorical, correlational, longitudinal, and non-parametric analyses without requiring estimation of marginal treatment effects.

The identification logic underlying DML, as formalized in Chapter 2, remains central to the conceptual framing of the study. In particular, the emphasis on separating structural relationships from confounding influences informs how results are interpreted, even in the absence of a full algorithmic implementation. As richer data on ex-ante AI ROI expectations and higher-resolution architectural complexity metrics become available, the analytical framework developed in this thesis provides a clear and coherent extension path for future causal estimation.

Conclusions and Recommendations

This Master's thesis set out to examine whether and how EA maturity influences the realization of AI-driven operational value in manufacturing organizations. Drawing on panel data from large manufacturing firms and employing a combination of descriptive, categorical, correlational, and ordinal-robust analytical techniques, the study provides convergent empirical evidence that EA maturity functions as a structural enabler of AI value creation rather than as a secondary or purely technical support capability.

Across multiple analytical dimensions, the results demonstrate a consistent pattern: firms with higher EA maturity are more likely to scale AI initiatives beyond pilots, to sustain performance gains over time, and to exhibit lower variance and downside risk in realized outcomes. These findings substantiate the central synthesis proposition (H5) that EA maturity operates as a value conversion capability, transforming AI potential into persistent, scalable, and predictable operational performance. In doing so, the study extends prior theoretical work on architecture-driven execution by providing empirical evidence in the context of AI-enabled manufacturing environments.

Importantly, the analysis also clarifies why AI adoption alone is an insufficient basis for competitive advantage. While localized AI initiatives can yield short-term improvements across a wide range of firms, sustained value creation depends on architectural coherence that supports integration, learning transfer, and repeatability. Without such foundations, organizations risk remaining trapped in prolonged experimentation cycles - commonly described as pilot purgatory - despite substantial investment in data and analytics capabilities.

Managerial and strategic recommendations

From a managerial perspective, the findings carry several concrete implications. First, executives should treat AI transformation not as a portfolio of isolated technology projects but as an architectural program that requires coordinated investment in processes, integration standards, and governance mechanisms. Incremental improvements in model performance or data volume are unlikely to yield durable returns unless they are embedded within a coherent EA.

This sequencing logic is consistent with the "foundation for execution" view of architecture-driven scalability (Ross et al., 2006) and with evidence that IT capability and complementary organizational practices are associated with persistent performance differences across firms (Bharadwaj, 2000). It is also compatible with research showing that value from analytics and data-driven decision-making emerges when organizations institutionalize processes and managerial routines rather than treating technology as a stand-alone tool (Brynjolfsson, Hitt, & Kim, 2011; Aral et al., 2012).

Second, investment sequencing matters. The evidence suggests that early and sustained investment in EA maturity lowers the cost and risk of future AI initiatives by enabling reuse and reducing outcome volatility. Firms that postpone architectural alignment until after large-scale AI deployment may achieve early wins but face significant barriers when attempting to scale or sustain those results.

Third, governance structures should reflect the architectural nature of AI value creation. Decision rights related to AI deployment, data ownership, and process change need to be explicitly defined and aligned with architectural standards. In this sense, EA maturity supports not only technical integration but also organizational clarity, improving the reliability of strategic and operational decision-making under uncertainty.

Methodological limitations

Several methodological limitations should be acknowledged to contextualize the interpretation of results. First, EA maturity is measured using observable architectural indicators derived from publicly disclosed information. While this approach enables consistent cross-firm comparison, it may not fully capture informal practices, undocumented governance routines, or tacit coordination mechanisms within organizations. Consequently, measured EA maturity should be interpreted as a proxy for architectural coherence, rather than as an exhaustive representation of all architectural activities.

Second, realized AI value is primarily captured through an ordinal operational index supported by disclosed qualitative and quantitative evidence. Explicitly disclosed financial AI ROI remains sparse across firms and time periods. As a result, operational value serves as the primary outcome variable, with financial ROI treated as supplementary where available. This focus reflects data realities rather than conceptual preferences and may underrepresent certain categories of benefits such as improved decision quality, risk avoidance, or organizational learning that are inherently difficult to monetize.

Third, although the study adopts a causal identification framework inspired by DML, the empirical implementation prioritizes transparent, hypothesis-consistent estimation over full algorithmic DML pipelines. Given the moderate sample size and the ordinal nature of the outcome variable, the analysis emphasizes robustness checks and structural interpretation rather than precise estimation of marginal treatment effects. Consequently, the results should be interpreted as evidence of structural relationships rather than as precise causal coefficients.

These limitations define clear empirical boundaries rather than weaknesses in the underlying research logic. All methodological choices are aligned with the nature of the available data and are reported transparently.

Implications for future research

The findings of this study open several promising avenues for future research. First, as richer data on ex-ante AI ROI expectations become available - through internal disclosures, surveys, or case-based research - future studies could explicitly quantify expectation–realization gaps and test more granular hypotheses related to organizational bias correction. Such data would enable direct estimation of behavioral effects that are only indirectly addressed in the present analysis.

Second, higher-resolution measures of architectural complexity, such as integration density, data latency variance, or system coupling metrics - would allow more precise modeling of architectural mechanisms. Combining such measures with longitudinal data would support more advanced causal designs, including full DML implementations.

Third, future research could explore heterogeneity across manufacturing sub-sectors, ownership structures, or regulatory environments to examine whether the architecture - AI value relationship varies systematically by context. Comparative case studies may also enrich understanding of how EA maturity evolves in practice and how firms successfully transition out of pilot-dominated AI strategies.

Finally, extending the analytical framework beyond manufacturing to other asset-intensive industries would test the generalizability of the architecture-driven value conversion logic proposed in this thesis.

List of References and Sources

Agrawal, A., Gans, J., & Goldfarb, A. (2018). *Prediction machines: The simple economics of artificial intelligence*. Harvard Business Review Press.

Agrawal, A. K., Gans, J. S., & Goldfarb, A. (2018). *Exploring the impact of artificial intelligence: Prediction versus judgment (NBER Working Paper No. 24626).* National Bureau of Economic Research. https://doi.org/10.3386/w24626

Aral, S., Brynjolfsson, E., & Wu, L. (2012). Three-way complementarities: Performance pay, human resource analytics, and information technology. *Management Science*, *58*(5), 913-931. https://doi.org/10.1287/mnsc.1110.1460

Athey, S., & Imbens, G. W. (2017). The state of applied econometrics: Causal inference with machine learning. *Journal of Economic Perspectives, 31*(2), 3-30. https://doi.org/10.1257/jep.31.2.3

Bachoo, S. (2018). Achieving value from enterprise architecture maturity. In *Proceedings of the Annual Conference of the South African Institute of Computer Scientists and Information Technologists (SAICSIT'18)* (pp. 286-295). Association for Computing Machinery. https://doi.org/10.1145/3278681.3278715

Baldwin, C. Y., & Clark, K. B. (2000). *Design rules, Vol. 1: The power of modularity.* MIT Press. https://doi.org/10.7551/mitpress/2366.001.0001

Baltagi, B. H. (2005). *Econometric analysis of panel data* (3rd ed.). John Wiley & Sons.

Bharadwaj, A. S. (2000). A resource-based perspective on information technology capability and firm performance: An empirical investigation. *MIS Quarterly, 24*(1), 169-196. https://doi.org/10.2307/3250983

Bizzdesign. (2025). *State of enterprise architecture*. https://bizzdesign.com/

Bresnahan, T. F., & Trajtenberg, M. (1995). General purpose technologies: Engines of growth? *Journal of Econometrics*, 65(1), 83–108. https://doi.org/10.1016/0304-4076(94)01598-T

Brynjolfsson, E., Hitt, L. M., & Kim, H. H. (2011). Strengthening the links between information technology and productivity: A longitudinal extension of earlier findings. *MIS Quarterly, 35*(3), 705-734.

Brynjolfsson, E., & Mitchell, T. (2017). What can machine learning do? Workforce implications. *Science, 358*(6370), 1530-1534. https://doi.org/10.1126/science.aap8062

Cai, C. (2025). *Agentic AI in manufacturing: 5 production-floor wins that deliver real ROI*. ChaiOne. https://www.chaione.com/blog/ai-manufacturing-operations-agentic-ai-production-floor-roi

Chernozhukov, V., et al. (2018). Double/debiased machine learning for treatment and structural parameters. *The Econometrics Journal, 21*(1), C1-C68. https://doi.org/10.1111/ectj.12097

Cohen, W. M., & Levinthal, D. A. (1990). Absorptive capacity: A new perspective on learning and innovation. *Administrative Science Quarterly, 35*(1), 128-152. https://doi.org/10.2307/2393553

Farbmacher, H., et al. (2022). Causal mediation analysis with double machine learning. *The Econometrics Journal*, *25*(2), 277–300. https://doi.org/10.1093/ectj/utac003

Foorthuis, R. M., et al. (2016). A theory building study of enterprise architecture practices and benefits. *Information Systems Frontiers, 18*(3), 541-564. https://doi.org/10.1007/S10796-014-9542-1

Hammer, M. (1990). Reengineering work: Don't automate, obliterate. *Harvard Business Review*, 68(4), 104–112.

Hazen, B. T., et al. (2017). Enterprise architecture: A competence-based approach to achieving agility and firm performance. *International Journal of Production Economics, 193*, 566-577. https://doi.org/10.1016/j.ijpe.2017.08.022

Hsiao, C. (2003). *Analysis of panel data* (2nd ed.). Cambridge University Press.

Iansiti, M., & Lakhani, K. R. (2020). *Competing in the age of AI: Strategy and leadership when algorithms and networks run the world*. Harvard Business Review Press.

Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus and Giroux.

Kotusev, S. (2017). Enterprise architecture: What did we study? *International Journal of Cooperative Information Systems, 26*(4). https://doi.org/10.1142/S0218843017300029

Kotusev, S. (2018). TOGAF-based enterprise architecture practice: An exploratory case study. *Communications of the Association for Information Systems, 43*, Article 20, 321-359. https://doi.org/10.17705/1CAIS.04320

LeanIX. (2019). *Enterprise architecture management insights report*. https://leanix.net/

Niemi, E., & Pekkola, S. (2020). The benefits of enterprise architecture in organizational transformation. *Business & Information Systems Engineering, 62* (6), 585-597. https://doi.org/10.1007/s12599-019-00605-3

National Institute of Standards and Technology & SEMATECH. (2022). *NIST/SEMATECH e-Handbook of statistical methods* (NIST Handbook 151). https://doi.org/10.18434/M32189

Ross, J. W., Weill, P., & Robertson, D. C. (2006). *Enterprise architecture as strategy: Creating a foundation for business execution*. Harvard Business School Press.

Saunders, M., Lewis, P., & Thornhill, A. (2016). *Research methods for business students* (7th ed.). Pearson Education.

Stahl, B. C., et al. (2023). Data or business first? Manufacturers' transformation toward data-driven business models. *Schmalenbach Journal of Business Research, 75*(3), 303–343. https://doi.org/10.1007/s41471-023-00154-2

Summary in English

Master's thesis: Architecting Competitive Advantage in Manufacturing: The Role of Enterprise Architecture Maturity in Realizing AI Value - a Causal Inference Analysis

Scope: 63 pages, 10 tables, 3 figures, 30 references.

This Master's thesis examines how Enterprise Architecture (EA) maturity influences the realization of artificial intelligence (AI) value in manufacturing organizations. While AI is widely recognized as a transformative technology, many firms struggle to translate AI initiatives from localized experiments into sustained operational and economic benefits. The aim of the thesis is to investigate whether EA maturity functions as a structural capability that enables AI initiatives to scale, persist, and deliver predictable value over time.

The research adopts a quantitative empirical design using panel data from large manufacturing firms. EA maturity is operationalized through observable architectural indicators reflecting process standardization, system integration, data coherence, and governance alignment. AI outcomes are measured primarily through an ordinal Operational AI Value Index capturing realized operational improvements, with financial ROI considered where publicly disclosed. The empirical analysis employs descriptive statistics, cross-tabulation with Chi-Square tests, correlation analysis, and ordinal-appropriate non-parametric robustness checks to evaluate structural relationships between EA maturity and AI outcomes.

The results provide consistent evidence that EA maturity is strongly associated with higher levels of realized AI value. Firms with higher EA maturity are significantly more likely to transition from pilot-level AI initiatives to scalable, enterprise-wide deployments. Moreover, AI-driven performance improvements in high-maturity firms persist over time and exhibit substantially lower variability compared to firms with fragmented architectural foundations. These findings indicate that EA maturity not only enhances AI performance but also reduces outcome volatility and operational risk.

Synthesizing the empirical results, the thesis concludes that competitive advantage does not arise from adoption of AI alone. Instead, EA maturity functions as a value-conversion capability that transforms AI potential into persistent, scalable, and predictable operational performance. This conclusion aligns with and extends the "foundation for execution" perspective in the Enterprise Architecture literature by providing empirical evidence in the context of AI-enabled manufacturing environments.

The study acknowledges methodological limitations related to the use of publicly observable architectural indicators and limited availability of disclosed financial AI ROI data.

Nevertheless, these constraints define transparent empirical boundaries rather than weaknesses of the research logic. The analytical framework developed in the thesis provides a clear extension path for future research using richer firm-level data and advanced causal estimation methods.

The results of the thesis are available in the full Master's Thesis submitted to Vilnius University, Faculty of Economics and Business Administration, and may be used for academic and educational purposes.

Summary in Lithuanian

Magistro darbas: Konkurencinio pranašumo kūrimas gamyboje: įmonių architektūros brandos vaidmuo realizuojant Dirbtinio Intelekto vertę – priežastinės analizės tyrimas

Apimtis: 63 puslapiai, 10 lentelių, 3 paveikslai, 30 šaltinių.

Šiame magistro darbe analizuojamas įmonių architektūros brandos (angl. Enterprise Architecture, EA) vaidmuo kuriant ir realizuojant dirbtinio intelekto (DI) vertę gamybos įmonėse. Nors DI technologijos plačiai pripažįstamos kaip transformuojančios verslo praktikas, daugelis organizacijų susiduria su sunkumais bandydamos perkelti DI iniciatyvas iš lokalių bandomųjų projektų į tvarią, organizacijos mastu kuriamą operacinę vertę. Darbo tikslas - empiriškai įvertinti, ar EA branda veikia kaip struktūrinis gebėjimas, leidžiantis DI sprendimus išplėsti, išlaikyti ilgalaikėje perspektyvoje ir užtikrinti prognozuojamus rezultatus.

Tyrimas grindžiamas kiekybiniu empiriniu dizainu, taikant tarptautinių gamybos įmonių panelinius duomenis. Įmonių architektūros branda vertinama remiantis stebimais architektūriniais požymiais, apimančiais procesų standartizaciją, sistemų tarpusavio integraciją, duomenų nuoseklumą ir valdymo struktūrų aiškumą. DI rezultatai vertinami naudojant operacinės DI vertės lygmenis aprašantį indeksą, papildomai atsižvelgiant į finansinę investicijų grąžą (ROI), kai tokia informacija yra viešai prieinama. Analizėje taikoma aprašomoji statistika, kryžminės analizės su χ^2 testais, koreliacinė analizė bei neparametriniai patikimumo testai, tinkami ranginio pobūdžio duomenims.

Empiriniai rezultatai atskleidžia stiprų teigiamą ryšį tarp EA brandos ir realizuotos DI vertės. Aukštos EA brandos įmonės reikšmingai dažniau pereina nuo bandomųjų DI iniciatyvų prie organizacijos mastu taikomų sprendimų, geba išlaikyti pasiektus rezultatus ilgalaikėje perspektyvoje ir patiria mažesnius rezultatų svyravimus. Tai leidžia daryti išvadą, kad EA branda ne tik didina DI teikiamą naudą, bet ir veikia kaip riziką mažinantis mechanizmas, prisidedantis prie stabilesnių bei labiau prognozuojamų rezultatų.

Apibendrinant tyrimo rezultatus, daroma išvada, kad konkurencinis pranašumas kyla ne iš pačių DI technologijų diegimo savaime, bet iš jų sąveikos su brandžia įmonių architektūra. EA branda veikia kaip vertės transformavimo mechanizmas, leidžiantis DI potencialą paversti tvaria, išplečiama ir prognozuojama operacine verte. Tyrimo rezultatai pateikiami visame magistro darbe, pateiktame Vilniaus universiteto Ekonomikos ir verslo administravimo fakultetui, ir yra prieinami akademiniais bei studijų tikslais.

Annexes

Appendix A. Descriptive Statistics

Table A.1 *Frequency distribution of the Operational AI Value Index*

Operational AI Value Index	Description	N	%	Cumulative %
0	No observable operational AI value	1	1.3	1.3
1	Local or pilot-level operational AI improvements	27	34.6	35.9
2	Scaled operational AI value (multi-site/process)	17	21.8	57.7
3	Systemic, enterprise-level AI value	33	42.3	100
Total		78	100	

Note. The Operational AI Value Index is an ordinal variable ranging from 0 (no observable operational AI value) to 3 (systemic, enterprise-level AI value). Frequencies are based on 78 valid firm-year observations.

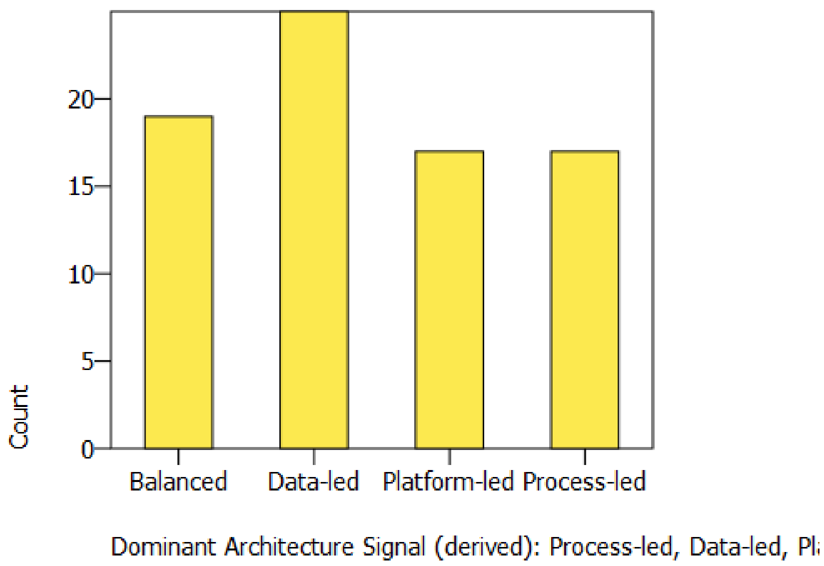
Table A.2 *Frequency distribution of Dominant Architecture Signal (N = 78)*

Dominant Architecture Signal	Description	N	%	Cumulative %
Balanced	No dominant architectural orientation	19	24.4	24.4
Data-led	Data and analytics-driven architecture focus	25	32.1	56.4

Platform-led	Platform or technology-centric architecture	17	21.8	78.2
Process-led	Process standardization-driven architecture	17	21.8	100
Total		**78**	**100**	

Note. Dominant Architecture Signal is a derived descriptive indicator identifying the primary architectural orientation (process-led, data-led, platform-led, or balanced) based on disclosed firm-year evidence. Percentages are based on valid observations only.

Figure A.2 *Distribution of Dominant Architecture Signal (N = 78)*



Note. The figure displays the distribution of the derived Dominant Architecture Signal categories across the sample. The indicator is used for descriptive profiling purposes only and is not applied as a causal regressor in hypothesis testing.

Appendix B. Crosstabulations and Chi-square Tests

Table B.1 *Crosstabulation of EA Maturity and Operational AI Value Index*

EA Maturity Level	**Operational AI Value = 0**	**1**	**2**	**3**	**Row Total**
EA = 3	1 (20.0%)	4 (80.0%)	0 (0.0%)	0 (0.0%)	5 (100%)
EA = 4	0 (0.0%)	23 (88.5%)	2 (7.7%)	1 (3.8%)	26 (100%)
EA = 5	0 (0.0%)	0 (0.0%)	15 (31.9%)	32 (68.1%)	47 (100%)
Column Total	**1**	**27**	**17**	**33**	**78**

Table B.2 *Pearson Chi-Square and Cramér's V*

Test	**Value**	**df**	**Asymptotic Sig. (2-sided)**
Pearson Chi-Square	79.88	6	< 0.001
Likelihood Ratio	88.29	6	< 0.001
Linear-by-Linear Association	52.78	1	< 0.001

Effect Size Measure	**Value**
Phi	1.01
Cramér's V	0.72

Table B.3 *EA_Maturity_Proxy × Pilot_vs_Scale*

EA Maturity Level	**Pilot (0)**	**Scaled (1)**	**Row Total**
EA = 3	1 (20.0%)	4 (80.0%)	5 (100%)
EA = 4	0 (0.0%)	26 (100.0%)	26 (100%)
EA = 5	0 (0.0%)	47 (100.0%)	47 (100%)
Column Total	**1**	**77**	**78**

Test	**Value**	**df**	**Asymptotic Sig. (2-sided)**
Pearson Chi-Square	14.79	2	0.001
Likelihood Ratio	5.7	2	0.058
Linear-by-Linear Association	6.28	1	0.012

Effect Size Measure	Value
Phi	**0.44**
Cramér's V	**0.44**

Note. The table reports the cross-tabulation between EA maturity levels and the Pilot_vs_Scale indicator based on 78 valid firm-year observations. Pilot_vs_Scale distinguishes between AI initiatives that remain at the pilot or localized stage (0) and those that have been deployed at scale across multiple sites or processes (1). Pearson's Chi-Square test is used to assess statistical independence. Effect size is reported using Cramér's V. All interpretations of results are provided in Section 3.3.

Appendix C. Correlation and Non-parametric Robustness Tests

Table C.1 *Correlation between EA maturity and Operational AI Value Index*

Variables	1	2
1. EA Maturity (EA_Maturity_Proxy)	1	0.828***
2. Operational AI Value Index	0.828***	1

N = 78

*****p < 0.001 (two-tailed)**

Note. Table C.1 reports Pearson correlation coefficients between EA maturity and the Operational AI Value Index based on 78 valid firm-year observations. Statistical significance is assessed using two-tailed tests. Interpretation of results is provided in Section 3.4.

Table C.2 *Kruskal–Walli's test of Operational AI Value Index across EA maturity levels*

EA Maturity Level	N	Mean Rank
EA = 3	5	12.2
EA = 4	26	18.5
EA = 5	47	54.02
Total	**78**	

Test Statistic	Value
Kruskal–Wallis H	56.03
df	2
Asymptotic Sig. (p-value)	< 0.001

Note. Table C.2 reports the results of a Kruskal-Wallis H test assessing whether distributions of the Operational AI Value Index differ across EA maturity levels. Mean ranks are reported for each maturity group. The test is based on 78 valid firm-year observations. Interpretation of results is provided in Section 3.4.