

Behavioral Factors as Determinants of Successful Scaling of Artificial Intelligence Pilot Projects in a Corporate Environment

Aneesha Sharma

Business Strategist, Chicago, USA

Abstract: *The article examines behavioral factors that shape the transition of corporate artificial intelligence (AI) initiatives from pilot deployments to scalable, sustained use. Relevance follows from the growing gap between rapid experimentation with AI tools and limited organizational capability to institutionalize them in everyday workflows. Novelty lies in integrating adoption frameworks (TAM and TOE) with evidence on human-AI interaction regarding trust calibration, cognitive load, and affective reactions, and in translating these constructs into a scaling-oriented conceptual framework. The paper aims to synthesize recent research and propose an analytical structure for diagnosing disengagement and “pilot-to-production” failure patterns, with special attention to AI coding assistants as a high-visibility class of corporate AI. Methods combine targeted literature synthesis, comparative conceptual analysis, and framework building. Recent scholarly and institutional sources are reviewed to derive constructs, hypothesized links, and governance implications. The paper concludes by articulating expected outcomes for management practice and a research agenda for future mixed-methods validation. The results inform leaders and designers of workplace AI.*

Keywords: AI scaling, AI capability, executive sponsorship, pilot-to-production, technology acceptance model, TOE framework, trust in AI, cognitive load, organizational influence, disengagement triggers, data readiness

1. Introduction

Corporate AI projects increasingly rely on pilots, yet many initiatives stall before enterprise-wide adoption, producing a recurring “pilot-to-production” gap whose behavioral roots remain under examined in managerial writing. Behavioral drivers matter because scaling requires repeated voluntary use, contextualization, and coordination across developers, managers, and front-line adopters, not only technical readiness. The purpose of the article is to systematize behavioral determinants of scaling success and to frame disengagement after initial uptake as a predictable failure mode in corporate AI tool adoption. The paper focuses on a salient contemporary case- AI coding assistants- because they concentrate trust, workload, and governance tensions in day-to-day knowledge work, while remaining representative of broader AI-enabled workflow change.

Recent literature suggests that the pilot-to-production gap is sustained not only by user-side hesitation but by a recurring structural pattern: pilots remain isolated, senior management backing weakens after experimentation, deployment capabilities lag behind model development, and organizations fail to connect use-case selection with integration, change management, and long-term ownership [10, 11, 15, 21]. Research on AI project failure reaches a similar conclusion by identifying problem misframing, weak data foundations, insufficient infrastructure, and organizational constraints as recurrent reasons why promising initiatives stall before operational scale [12, 13]. For this reason, AI scaling should be examined as a joint behavioral and organizational process rather than as a post-adoption extension of initial acceptance alone.

The objective is to synthesize recent scholarly evidence using the Technology Acceptance Model and the Technology–Organization–Environment framework, extending them with

constructs from trust research in AI and human–AI interaction studies. Three tasks guide the article:

- 1) To identify behavioral constructs that condition scaling beyond initial pilot acceptance;
- 2) To map how usefulness perceptions, trust calibration, user challenges, organizational influence, and disengagement triggers interact during scale-up;
- 3) To derive practical implications for governance and change management that can be tested in future empirical work.

The article advances a scaling-centered conceptual model that links behavioral adoption dynamics to organizational readiness, operationalization capability, and governance conditions, thereby clarifying why some AI pilots evolve into routinized enterprise systems while others remain temporary experiments or fail during transition to production.

2. Materials and Methods

The analytical base integrates 21 peer-reviewed, scholarly, or institutional sources, primarily published within the last five years, with the inclusion of one earlier study due to its direct relevance to scaling data-driven pilot projects. S. Afroogh [1] conceptualizes trust/distrust in AI and differentiates trust-breakers and trustmakers that shape acceptance trajectories. F. Ali [2] empirically extends TAM for AI coding assistants by operationalizing trust-related variables that influence perceived usefulness and intentions. S. J. Daly [3] examines organizational AI adoption through shifting attitudes and the calibration of trust across organizational roles. L. Hughes [4] applies the TOE lens to generative AI adoption and links complexity, skills, and organizational capacity to adoption outcomes. R. Lucas [5] develops a calibrated-trust perspective and provides actionable framing for evaluating trustworthiness evidence in human–AI collaboration. NIST [6] provides a risk-management structure for trustworthy AI

that supports governance of scaling decisions through harm-aware controls. A. S. Pinto [7] meta-analyzes TOE determinants of organizational AI adoption across industries, supporting cross-context factor generalization. M. Taeihagh [8] synthesizes governance challenges for generative AI, clarifying policy-relevant levers in organizational settings. A. Valtonen [9] empirically connects AI adoption conditions to employee well-being, informing behavioral sustainability during scale-up. M. Berndtsson, A. Ericsson, and T. Svahn [10] examine barriers to scaling data-driven pilot projects and highlight the recurrent problem of pilots remaining siloed with limited business value, especially under weak senior-management support. J. Vallone [11] reassesses widely cited AI failure claims and synthesizes evidence that pilot-to-production gaps are more strongly associated with organizational constraints such as problem framing, data readiness, sponsorship, integration, and change management than with purely technical limitations. D. Schlegel [12] identifies failure factors of AI projects through expert interviews and shows that obstacles arise across technical, organizational, and human dimensions. J. Ryseff, B. F. De Bruhl, and S. J. Newberry [13] analyze root causes of AI project failure and emphasize recurring anti-patterns related to poor problem definition, insufficient infrastructure, and weak data foundations. J. Kutz, J. Neuhüttler, J. Spilski, and T. Lachmann [14] examine implementation of AI technologies in manufacturing and identify success factors and challenges related to data, development and operational processes, and stakeholder engagement. D. Leff and K. T. K. Lim [15] discuss the organizational barriers that prevent firms from moving beyond experimentation, including fragmented enterprise data, legacy systems, and limited AI-ops capabilities. M. I. Merhi [16] evaluates critical success factors affecting AI implementation and supports a broader reading of organizational, technological, process, and environmental conditions. P. Mikalef and M. Gupta [17] conceptualize AI capability at the firm level and connect it to organizational outcomes, helping explain why scaling depends on more than local pilot success. S. O. Al Hleewa and M. Al Mubarak [18] summarize success factors of AI use and support the inclusion of technological, organizational, and environmental conditions in the analysis of implementation outcomes. J. A. Meyer [19] discusses success factors and development areas for implementing generative AI in companies, extending the analysis from pilot experimentation to company-level deployment conditions. N. Sambasivan, S. Kapania, H. Highfill, et al. [20] introduce the concept of data cascades and show how undervalued data work produces delayed downstream failures that can undermine AI deployment at scale. S. Makinani and M. B. Nagaraja [21] propose a multi-stage roadmap for scaling AI from project pilots to program-wide transformations by linking technical readiness, executive sponsorship, MLOps operationalization, business alignment, and change management.

For writing the article, a structured literature synthesis, a comparative conceptual analysis across the TAM/TOE and trust frameworks, and the construction of an inductive–deductive framework were applied, with a governance-oriented interpretation aligned with responsible AI scaling guidance.

3. Results

Evidence from recent trust-in-AI scholarship supports treating sustained use as a regulated behavioral state rather than a simple extension of initial acceptance: trust and distrust function as adoption “controls” that can accelerate diffusion or suppress it, and they can be destabilized by perceived threats to autonomy, dignity, or fairness, which are directly relevant when AI tools intervene in expert judgment and professional identity [1]. In scaling contexts, this implies that pilots that “work” technically may still fail behaviorally if they trigger persistent doubt, social risk, or reputational anxiety for the user group expected to operationalize them.

A TAM-consistent reading of AI tool uptake remains helpful in scaling, but only when interpreted dynamically. When perceived usefulness is tied to locally verifiable gains—time savings, fewer workflow interruptions, higher output quality—behavioral intention can stabilize into routine use; when usefulness is ambiguous or costly to verify, users fall back to heuristics such as peer narratives, managerial signaling, and prior experiences with automation. The extended TAM study of AI coding assistants operationalizes this mechanism by showing that trust-related beliefs influence perceived usefulness and adoption intentions. At the same time, negative affective variables (e.g., insecurity or doubt) can attenuate acceptance pathways, even in settings where enthusiasm for AI tooling is high [2]. For corporate scaling, the implication is that “usefulness” cannot be treated as a static attribute of the tool; it is produced through organizational instrumentation (measurement, feedback, integration) and user learning that reduce uncertainty and interaction costs.

TOE-based evidence complements TAM by clarifying why behavioral readiness is constrained by organizational capacity and environmental conditions. In generative AI adoption research using the TOE, complexity and skill constraints appear alongside change capacity and organizational readiness as determinants of whether AI initiatives move beyond experimentation into routinized processes [4]. The TOE meta-analysis reinforces the cross-industry relevance of technological, managerial, and environmental determinants and justifies treating scale-up failure as a multi-level phenomenon: individual acceptance can coexist with organizational incapacity to institutionalize tool-supported work, creating churn in use patterns and intermittent disengagement [7]. In corporate AI pilots, this translates into a common failure pathway: early adopters achieve localized success, but diffusion stalls when onboarding, process ownership, and governance are not allocated, leading the broader population to view the tool as “optional,” risky, or disruptive.

Yet the literature on pilot-to-production transitions suggests that organizational readiness should be specified more concretely than in generic TOE terms. Studies of AI pilots and failed AI projects repeatedly identify executive sponsorship, data readiness, integration capacity, deployment infrastructure, and change-management continuity as the practical conditions that determine whether local proof-of-concept value becomes enterprise value [10, 11, 13, 15, 21]. In this reading, scale-up failure does not arise simply because

users stop trusting the tool; it often begins when the organization cannot preserve ownership, funding, technical support, and workflow fit after the pilot team exits the experimentation phase.

Trust calibration emerges as a central behavioral bridge between pilot enthusiasm and scaling durability. Organizational adoption research indicates that trust varies by role- developers, managers, and users evaluate AI differently- and that positive attitudes can evolve into calibrated trust grounded in realistic views of AI limitations when organizational support and visible benefits are present [3]. This pattern is critical for AI coding assistants in corporate environments, where developers may value speed and ideation support, while managers focus on predictability, auditability, and risk. Calibrated trust is not synonymous with high trust; it is a fit between user expectations and system performance bounds. A calibrated-trust approach operationalizes this by emphasizing evidence and transparency in how trustworthiness is communicated, so that users can align their reliance decisions with the tool's actual competence envelope [5]. In scaling terms, teams that reach calibrated trust can adopt explicit "reliance policies" (when to use, when to verify, when to avoid), reducing emotional friction and preventing the oscillation between overreliance and rejection that frequently follows pilot hype.

Cognitive load and interaction cost are additional behavioral determinants of scale. Even when usefulness is high, sustained use can degrade if the tool introduces mental overhead: prompt crafting, output validation, context switching, or rework caused by subtle errors. In practice, scaling succeeds when interaction costs are progressively reduced through better integration into IDEs and pipelines, standardized prompting practices, and shared artifacts that support verification. Trust and cognitive load are coupled: when users anticipate frequent verification, perceived effort rises; when verification is supported by process and tooling, perceived effort becomes manageable and reliance becomes stable [2]. This coupling explains why developer disengagement after early adoption can occur without any dramatic performance collapse: a tool that feels cognitively expensive in real tasks can be abandoned in favor of familiar workflows, especially under deadline pressure.

A related but analytically distinct source of scaling failure lies in the invisibility of data work. Organizations often invest symbolic attention in models while underinvesting in data preparation, maintenance, lineage, accessibility, and production-grade monitoring. Research on data cascades shows that unresolved data issues generate delayed downstream effects that remain hidden during early experimentation but intensify during broader deployment [20]. This observation aligns with AI failure studies that identify weak data foundations and inadequate infrastructure among the most recurring reasons why AI initiatives do not survive the transition from demonstration to operational use [13, 15]. For corporate AI pilots, data readiness is therefore not a background technical condition; it is a scaling determinant that shapes reliability perceptions, rework intensity, and the credibility of the tool in everyday workflows.

Organizational influence functions as a behavioral "field" that shapes whether pilot outcomes are interpreted as credible and whether usage becomes socially safe. Evidence on governance and organizational adoption indicates that scaling requires explicit responsibility allocation, policy guardrails, and a shared understanding of acceptable use, particularly for generative AI [8]. When these conditions are absent, developers may perceive adoption as personally risky (due to security, compliance, and code quality liabilities), and managers may respond with ambiguous or inconsistent signals, amplifying disengagement triggers. Conversely, when leadership provides consistent framing- what is allowed, how to document use, how to review outputs- users can externalize part of the risk to the organization, lowering emotional and reputational barriers to use [8].

Beyond policy clarity and social legitimacy, organizations differ in their underlying AI capability. Prior research suggests that AI outcomes depend on a combination of firm-level resources, routines, and complementary competences rather than on isolated tools alone [17]. Success-factor and maturity-oriented studies reach a similar conclusion by emphasizing multi-domain readiness across technology, data, strategy, people, organizational arrangements, and process design [16, 18, 19]. This means that pilot success cannot be extrapolated directly into scale-up readiness: a team may report positive local experiences with an AI system while the wider organization still lacks the capability profile required for repeatable deployment, maintenance, and cross-unit diffusion.

To integrate these findings into a scaling-relevant evidence structure, Figure 1 anchors behavioral determinants to a harm-aware governance framing. Figure 1 is adapted from NIST's categorization of potential harms (to people, organizations, ecosystems), which is directly applicable to corporate scaling because perceived harm probability influences trust, reliance, and willingness to routinize use.

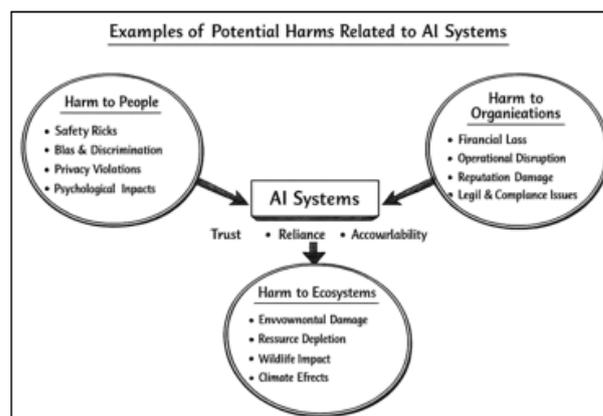


Figure 1: Examples of potential harms related to AI systems (developed by the author based on NIST AI RMF 1.0)

In a pilot phase, harms are often discounted as hypothetical, leading perceived usefulness to dominate. During scaling, the perceived harm surface expands: more users, more data touchpoints, broader operational exposure. Users then re-evaluate their reliance through the lens of organizational harm (e.g., breaches, monetary losses, reputational damage), which can trigger disengagement even if the tool remains personally productive. The behavioral determinant here is not only

objective risk but perceived accountability. If users believe errors will be attributed to them rather than treated as managed organizational risk, disengagement becomes a rational self-protective response.

Behavioral sustainability further depends on well-being and psychological conditions of work. Empirical evidence links AI use conditions to employee wellbeing, implying that scale-up policies that intensify surveillance, reduce autonomy, or increase workload volatility can erode engagement and provoke resistance or silent abandonment [9]. For developer-facing AI tools, this suggests that the scaling strategy must align productivity narratives with autonomy-preserving practices; otherwise, a tool introduced as “assistance” can be reinterpreted as managerial pressure, intensifying emotional fatigue and decreasing long-term use.

Synthesizing across the initial and newly added literature, the framework can be reformulated as a two-level scaling model. The first level captures behavioral mediators of sustained use: perceived usefulness, trust calibration, user-side challenges, organizational influence, and disengagement triggers. The second level captures organizational-enabling conditions that determine whether pilot value can be converted into routinized enterprise deployment: executive sponsorship, data readiness, operationalization capability, business-use-case fit, and AI capability/maturity [10–21]. Behavioral factors explain whether users continue to rely on the system; organizational-enabling conditions explain whether such reliance can be supported, stabilized, and reproduced beyond the pilot setting.

4. Discussion

The synthesized results indicate that scaling failures frequently reflect behavioral breakdowns arising from mismatches between individual-level acceptance and organization-level readiness. A practical implication is that corporate leaders should treat scaling as a behavioral systems problem: users do not “adopt the pilot”; they adopt a new routine under uncertainty, accountability, and social evaluation. Trust research clarifies that distrust is not merely a lack of information; threats to autonomy and dignity can trigger, and these triggers are amplified when organizational messaging frames AI as a replacement rather than an augmentation [1].

The added literature suggests that behavioral explanations are necessary but insufficient for a robust account of enterprise AI scaling. A pilot may achieve favorable user reactions and still fail to scale if the organization lacks executive sponsorship, reliable data pipelines, integration routines, deployment infrastructure, or a clearly owned business case [10, 11, 13, 15, 21]. Conversely, strong infrastructure and governance do not guarantee scale if users perceive the system as cognitively expensive, unsafe, or misaligned with professional autonomy. The most persuasive interpretation is therefore interactional: behavioral acceptance and organizational-operational readiness reinforce one another, while weakness in either layer can interrupt scale-up.

To translate the proposed scaling model into operational guidance for future study design and managerial diagnostics, Table 1 maps each construct to measurable indicators suitable for scaling corporate AI tools, with a focus on coding assistants as a representative class.

Table 1: Construct-to-indicator mapping for analyzing AI pilot scaling and disengagement (developed by the author based on [1–5, 10–21])

Core construct	Scaling-relevant interpretation	Example indicators for proposed measurement
Usefulness	Perceived net performance gain under real workflow constraints	self-reported productivity gain; perceived output quality; task completion time perception; intention to continue use
Trust	Reliance willingness is conditioned by perceived competence and integrity	reliance frequency; verification intensity; perceived reliability; perceived safety/ethical acceptability
Challenges	Friction sources that raise interaction cost and create avoidance	cognitive load proxies (effort, interruptions); error correction effort; emotional strain (insecurity/doubt); tool–workflow mismatch
Organizational influence	Norms, resources, and accountability framing shape perceived safety	clarity of allowed use; availability of training; managerial encouragement; review norms; governance maturity
Executive sponsorship	Continuity of leadership backing across post-pilot transition	identified executive sponsor; scale-up budget continuity; escalation speed; leadership review cadence
Data readiness	Availability and quality of production-grade data for stable deployment	data coverage; data accessibility; quality incidents; lineage clarity; retraining readiness
Operationalization capability	Capacity to deploy, integrate, monitor, and maintain AI systems in live workflows	MLOps/AIOps routines; integration depth; monitoring practices; rollback procedures; deployment frequency
Business-use-case fit	Degree of alignment between AI functionality, workflow pain point, and measurable business outcome	problem-definition clarity; KPI traceability; workflow compatibility; ownership of business case
Disengagement triggers	Events or conditions converting friction into reduced use/abandonment	perceived incident risk; negative peer feedback; policy ambiguity; workload spikes; wellbeing deterioration signals

A mixed-methods design can treat these indicators as survey constructs and qualitative codes, enabling triangulation between self-reported disengagement narratives and the organizational conditions that precede abandonment. TOE evidence strengthens the feasibility by providing an

established framework for modeling organizational readiness and environmental constraints that can be operationalized alongside TAM measures [4].

A second implication concerns governance interventions that

reduce behavioral friction without overstating the performance of the tools. NIST's harm framing clarifies why risk-aware communication and process controls can increase sustained use: users disengage when they perceive

unmanaged harm pathways and personal liability. Table 2 consolidates the governance and change levers supported by the reviewed sources and shows how each lever targets one or more constructs.

Table 2: Governance and change-management levers aligned to the proposed scaling model [4, 5, 8–21]

Lever	Mechanism of action	Primary construct(s) affected
Harm-aware policy and accountability framing	reduces perceived personal liability; clarifies acceptable reliance and escalation paths	trust; disengagement triggers; organizational influence
Training that teaches verification routines	lowers cognitive load by standardizing “use + check” practices; supports calibrated trust	trust; challenges
Integration into workflow and review pipelines	reduces context switching and rework; makes usefulness visible through artifacts	usefulness; challenges
Leadership signaling and resource allocation	stabilizes norms, legitimizes time for learning, and prevents “optional tool” interpretation	organizational influence; usefulness
Wellbeing-sensitive deployment practices	prevents adoption from being reframed as surveillance/pressure; protects autonomy	challenges; disengagement triggers
Executive sponsorship and portfolio governance	protects scale-up resources, legitimizes post-pilot ownership, and prevents pilot abandonment after initial enthusiasm declines	executive sponsorship; organizational influence; disengagement triggers
Data and deployment readiness programs	reduce hidden failure sources by improving data quality, infrastructure, integration, and monitoring before broad rollout	data readiness; operationalization capability; challenges

Table 2 implies a managerial caution: scaling that relies on pressure rather than capability building can degrade well-being and provoke disengagement that is rational from the employee perspective, particularly in high-autonomy professions such as software development [9]. In proposal-stage work, these levers can be treated as candidate moderators in a future model. For example, policy clarity and training intensity may weaken the link between challenges and disengagement triggers, while organizational readiness may strengthen the link between usefulness perceptions and sustained use.

For managerial diagnosis, this expanded model implies that AI pilots should be evaluated along a staged escalation logic rather than a single success/failure binary. Early-stage assessment should focus on problem framing and business-use-case fit; transition-stage assessment should address data readiness, sponsorship continuity, and deployment capability; scale-stage assessment should examine routinization, workflow integration, and sustained behavioral reliance [11, 13, 21]. Such sequencing is particularly relevant for generative AI, where rapid experimentation often creates an illusion of readiness that is not matched by enterprise integration, governance maturity, or repeatable operating routines.

5. Conclusion

The article argues that successful scaling of AI pilots depends on the interaction between two layers of conditions. AI coding assistants serve in this article as a high-visibility corporate use case through which these broader scaling dynamics become analytically observable. The first layer is behavioral and concerns usefulness perceptions, calibrated trust, user-side challenges, organizational influence, and disengagement triggers. The second layer is organizational-operational and concerns executive sponsorship, data readiness, business-use-case fit, operationalization capability, and broader AI capability within the firm. AI pilots tend to stall when one of these layers remains underdeveloped: positive pilot reactions do not translate into scale without ownership, infrastructure, and workflow integration, while strong technical preparation

does not translate into durable use without trust, manageable cognitive effort, and socially legitimate reliance. The proposed conceptual framework therefore repositions the pilot-to-production gap as a combined behavioral, organizational, and operational phenomenon and offers a basis for future mixed-methods testing in enterprise environments.

References

- [1] Afroogh, S., Akbari, A., Malone, E., et al. (2024). Trust in AI: Progress, challenges, and future directions. *Humanities and Social Sciences Communications*, 11, Article 1568. <https://doi.org/10.1057/s41599-024-04044-8>
- [2] Ali, F., Ahmed, A., Alipour, M. A., et al. (2025). Adoption of AI-coding assistants in programming education: Exploring trust and learning motivation through an extended technology acceptance model. *Journal of Computer Education*. Advance online publication. <https://doi.org/10.1007/s40692-025-00375-w>
- [3] Daly, S. J., Wiewiora, A., & Hearn, G. (2025). Shifting attitudes and trust in AI: Influences on organizational AI adoption. *Technological Forecasting and Social Change*, 215, 124108. <https://doi.org/10.1016/j.techfore.2025.124108>
- [4] Hughes, L., Davies, F., Li, K., Madugoda Gunaratne, S., Malik, T., & Dwivedi, Y. K. (2026). Beyond the hype: Organisational adoption of generative AI through the lens of the TOE framework—A mixed methods perspective. *International Journal of Information Management*, 86, 102982. <https://doi.org/10.1016/j.ijinfomgt.2025.102982>
- [5] Lucas, G. M., Becerik-Gerber, B., & Roll, S. C. (2024). Calibrating workers' trust in intelligent automated systems. *Patterns*, 5(9), 101045. <https://doi.org/10.1016/j.patter.2024.101045>
- [6] National Institute of Standards and Technology. (2023). *Artificial intelligence risk management framework (AI RMF 1.0)* (NIST AI 100-1). <https://www.nist.gov/itl/ai-risk-management-framework>

- [7] Pinto, A. S., Abreu, A., Pérez Cota, M., et al. (2025). A meta-analysis of TOE factors driving organizational adoption of artificial intelligence across industries. *Discover Artificial Intelligence*. Advance online publication. <https://doi.org/10.1007/s44163-025-00747-2>
- [8] Taciagh, A. (2025). Governance of generative AI. *Policy and Society*, 44(1), 1–22. <https://doi.org/10.1093/polsoc/puaf001>
- [9] Valtonen, A., Saunila, M., Ukko, J., Treves, L., & Ritala, P. (2025). AI and employee well-being in the workplace: An empirical study. *Journal of Business Research*, 199, 115584. <https://doi.org/10.1016/j.jbusres.2025.115584>
- [10] Berndtsson, M., Ericsson, A., & Svahn, T. (2020). Scaling up data-driven pilot projects. *AI Magazine*, 41(3), 94–102. <https://doi.org/10.1609/aimag.v41i3.5307>
- [11] Vallone, J. (2025). *Reassessing AI pilot failure rates: A scoping review and managerial implications*. SSRN Working Paper. <https://doi.org/10.2139/ssrn.5459054>
- [12] Schlegel, D., Schuler, K., & Westerberger, J. (2023). Failure factors of AI projects: Results from expert interviews. *International Journal of Information Systems and Project Management*, 11(3), Article 3. <https://aisel.aisnet.org/ijispm/vol11/iss3/3>
- [13] Ryseff, J., De Bruhl, B. F., & Newberry, S. J. (2024). *The root causes of failure for artificial intelligence projects and how they can succeed: Avoiding the anti-patterns of AI*. RAND Corporation. https://www.rand.org/pubs/research_reports/RRA2680-1.html
- [14] Kutz, J., Neuhüttler, J., Spilski, J., & Lachmann, T. (2022). Implementation of AI technologies in manufacturing – Success factors and challenges. In C. Leitner, W. Ganz, C. Bassano, & D. Satterfield (Eds.), *The human side of service engineering* (AHFE Open Access, Vol. 62). AHFE International. <https://doi.org/10.54941/ahfe1002565>
- [15] Leff, D., & Lim, K. T. K. (2021). The key to leveraging AI at scale. *Journal of Revenue and Pricing Management*, 20, 376–380. <https://doi.org/10.1057/s41272-021-00320-3>
- [16] Merhi, M. I. (2023). An evaluation of the critical success factors impacting artificial intelligence implementation. *International Journal of Information Management*, 69, 102545. <https://doi.org/10.1016/j.ijinfomgt.2022.102545>
- [17] Mikalef, P., & Gupta, M. (2021). Artificial intelligence capability: Conceptualization, measurement calibration, and empirical study on its impact on organizational creativity and firm performance. *Information & Management*, 58(3), 103434. <https://doi.org/10.1016/j.im.2021.103434>
- [18] Al Hleewa, S. O., & Al Mubarak, M. (2023). Success factors of using artificial intelligence. In M. Al Mubarak & A. Hamdan (Eds.), *Technological sustainability and business competitive advantage* (Internet of Things). Springer. https://doi.org/10.1007/978-3-031-35525-7_11
- [19] Meyer, J. A. (2025). Success factors and development areas for the implementation of generative AI in companies. *Junior Management Science*, 10(1). <https://doi.org/10.5282/jums/v10i1pp1-23>
- [20] Sambasivan, N., Kapania, S., Highfill, H., Akrong, D., Paritosh, P., & Aroyo, L. M. (2021). “Everyone wants to do the model work, not the data work”: Data cascades in high-stakes AI. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (CHI '21) (Article 39, pp. 1–15). Association for Computing Machinery. <https://doi.org/10.1145/3411764.3445518>
- [21] Makinani, S., & Nagaraja, M. B. (2025). Scaling AI from project pilots to program-wide transformations. *IJERET*, 6(3), 41–46. <https://ijeret.org/index.php/ijeret/article/view/242>