International Journal of Emerging Research in Engineering and Technology
Pearl Blue Research Group| Volume 6 Issue 3 PP 41-46, 2025
ISSN: 3050-922X | https://doi.org/10.63282/3050-922X.IJERET-V6I3P105

Original Article

Scaling AI from Project Pilots to Program-Wide Transformations

Shreya Makinani[1], Madhusudan Bangalore Nagaraja[2]
[1]Information and Systems Engineering, University of Southern California, California, USA.
[2]Technical Delivery Manager, Irving, TX, USA.

Received On: 24/05/2025 Revised On: 14/06/2025 Accepted On: 26/06/2025 Published On: 13/07/2025

Abstract *- Companies are spending billions on AI, but most initiatives get stuck in the pilot phase. This article presents the AI Scaling Navigator, a six-step framework that integrates technical, organizational, and managerial readiness into one actionable roadmap. Based on industry benchmarking, current literature, and cross-industry case syntheses, the Navigator maps the journey from use-case discovery to enterprise deployment in six stages Pilot Discovery; Data & Talent Readiness; Executive Sponsorship; MLOps Operationalization; Business Alignment & Change Management; and Scalable Deployment & Optimization. The framework requires aligning infrastructure, governance, and culture to convert experimentation into sustainable business value. Applied to retail, manufacturing, and financial services contexts, the Navigator is associated with higher deployment success, improved operational performance, and higher innovation potential. The article offers practical guidance to AI managers and transformational leaders who seek to scale AI responsibly and reproducibly across industries.*

Keywords *- AI adoption, AI scaling, change management, machine learning operations, organizational transformation, project management.*

1. Introduction

Artificial intelligence has been the primary business strategy enabler, with the potential to drive efficiency, innovation, and competitiveness in a big way. Yet, after massive investments, the jump from the successful pilot projects of AI to program-scale transformation continues to be plagued with challenges. Surveys suggest that over 70% [10] of the AI pilots do not scale to production, thus limiting the possibility to reengineer the operations and decision making.

Closing the perennial “pilot-to-production” gap to deployment of AI entails more than technological solutionsit entails mature organizational dynamics understanding along with successful change management expertise Eliminating these roadblocks entails the exploitation of the formal, evidence-based processes capable of transferring viable AI pilots to enterprise-wide traction. For this, this paper puts forth the AI Scaling Navigator: a proved six-step blueprint enabling organizations to take step-by-step jumps from the identification of initial use-case to organization-scale deployment. By having the Pilot Discovery, Data & Talent Readiness, Executive Sponsorship, Operationalization (MLOps), Business Alignment & Change Management, and Scalable Deployment & Optimization phases, the phases of the Navigator confront major roadblocks identified across cross-industry studies forthright. The success of the model comes to light yet further by virtue of being tested within-the-trenches across organizational settings diverse, detailed across subsequent sections of this paper. By and large, this blueprint avails organizations a stable, actionable roadmap to rooting out experimental AI pilots to lasting organizational benefits [9].

2. Methods

This paper utilized a multi-method strategy comprising:

- Benchmark Study Gathering failure-rate data of the leading institutions (IDC, Gartner, McKinsey, S&P Global) to identify systemic causes of AI pilot drop-out rates.
- Literature Review Scanning through more than 30 peer-reviewed papers and business whitepapers published between 2018–2025 to distill verifiable scaling techniques.
- Case Study Synthesis Picking anonymized case studies from retail, manufacturing, and financial services companies, by availability of quantifiable KPIs, scope of operating context, and well-documented problems to scale.
- Thematic analysis Plotting outcomes onto current AI project maturity models to establish the six phases of the AI Scaling Navigator.

This mixed-methods strategy guaranteed that the framework aims to balance the theoretical soundness with workability, with empirical findings in the sectors to support it.

3. AI Pilot Projects: Success Rates and Challenges

Studies and studies on the information industry commonly cite the failures of AI pilot projects in the 70–90% range [14]. Even established companies will shelf most of the pilots by murky ROI, under-resourcing, or scalability problems [11]. Most characteristic hurdles are below:

- Unclear Objectives/Return on Investment: Projects start without any specific business problem or measurable KPIs in view and thus cannot readily demonstrate tangible value [14].
- Poor Data and Talent: Insufficient amount of 'AI-ready' data and trained talent typically hold up development work. Validating audits and information quality and governance are the most commonly observed reasons for pilot failures in the technology industry.
- Poor Integration: Even with successful pilots, production rollouts have uncovered workflow, infrastructural, and legacy system integration issues.
- Inability to Adapt to Change: Poor prioritization and end-user resistance can lead to project slowdown beyond pilot level.
- Model Performance Degradation: After deployment, performance degrades for 91% of the models because of data drift and changing business context and must be continually monitored and retrained [16].

4. Sustaining AI Projects in Production (MLOps)

Deployment to production is not the end goal but in reality, continuous AI value creation requires strong MLOps [7] best practices:

- Continuous Monitoring: Prediction performance, data drift, and business outcomes must be tracked in real-time.
- Regular Retraining: Models must be retrained on a regular basis to offset deterioration over time.
- Robust Data Pipelines: Reliable, secure, and scalable data infrastructure supports model longevity.
- Governance & Traceability: Regulatory goals and lifecycle management must be supported by version control and clear ownership to facilitate regulatory objectives and model lifecycle management [7], [8].

5. The Central Role of Data Readiness

Data is the basis of all AI success. Key things:

- Volume and Diversity: Pilots need to scale with datasets that capture the complex nature of operating environments.
- Quality and Consistency: Data inaccuracies and inconsistencies are the main reasons why pilots fail.
- Accessibility and Integration: Enterprise-wide scaling is hindered by data silos and infrastructure that is fragmented.
- Privacy & Governance: Early understanding of compliance needs avoids legal and ethical obstructions.

As per a report by 2025 Gartner, 63% of the organizations were seen to lack AI-capable data pipelines and platforms, which put their projects in risk [15].

6. Literature Gaps and Best-Practice Recommendations

AI adoption in organizations has gained momentum over the last two years, but literature and practice reveal several unresolved gaps that prevent successful upscaling of AI initiatives. Overcoming these areas is necessary for the shift from isolated pilot projects to concrete, long-term AI programs.

6.1. ROI Measurement and Value Mapping

An ongoing gap is the absence of standard frameworks that quantitatively link technical model metrics (e.g., accuracy, F1-score) to concrete business outcomes [3] (e.g., cost savings, revenue growth, or customer satisfaction). Most organizations and case studies either report technical success or anecdotal business value, but few rigorously trace AI performance improvements to financial return measured or operational KPIs. This gap reduces stakeholder buy-in and makes AI resource planning at scale more difficult. Cost-benefit and value realization frameworks closing this translation gap are in urgent need to assist organizations in predicting, measuring, and reporting AI business value during the lifecycle [3].

6.2. Organizational Readiness and Maturity Models

Despite increasing interest, comprehensive and empirically validated maturity models for AI adoption are nascent [17]. Current IT maturity models do not address AI-specific challenges such as iterative development, data dependence, and ongoing maintenance needs [17]. Initial frameworks exist but require additional empirical validation and tailoring to organization sizes and sectors. An effective readiness model needs to assess technical, data, cultural, governance, and talent dimensions with actionable road maps for organizations before scaling pilots to production environments.

6.3. Early Integration of Data Governance and Ethics

We note in the literature that data governance, regulatory compliance, and ethics are typically left to later phases in the AI project life cycle. This may result in last-minute delays, non-compliance risk, or models that are subsequently found to be inappropriate to deploy. To mitigate these risks, best practice now demands the introduction of "ethical review gates" and stringent data governance checks at every milestone from ideation to continuous operation [13]. Proactively designing for fairness, explainability, privacy, and accountability strengthens stakeholder trust and compliance with emerging regulatory demands, e.g., GDPR or the EU AI Act [6].

6.4. The AI 'Translator' and Hybrid Talent Gap

Both domain/business expertise and technical AI literacy are required in project managers and leaders, a scarcity that continues to be a bottleneck. The majority of organizations lack anyone who can translate between data science teams and business stakeholders, slowing adoption and perplexing increasing misalignment between AI outputs and business goals. Best practice and literature dictate the establishment of specialist 'AI translator' roles through recruitment or internal up skilling to translate across the gap, maintain project relevance, and accelerate scaling. This can include AI-aware product managers, technically oriented business analysts, or rotational development schemes across the gap [9].

6.5. Continuous Monitoring and Feedback Loops

There is limited research and practice literature on long-term, comprehensive monitoring of AI system performance across both business value capture and technical health (e.g., model drift, data quality). Most published work focuses on initial deployment or short-term results, with little long-term business/KPI feedback and model retraining approaches in evolving, real-world environments. Developing robust feedback loops connecting model diagnostics and business impact metrics is a scholarship and practice frontier.

Table 1: Comparison of Traditional vs. AI Project Management

Criteria	Traditional Software Projects	AI-Based Projects
Development	Explicit programming	Learn from data (models)
Outcome	Deterministic	Probabilistic
Requirements	Upfront, fixed	Evolving, refined during pilots
Project Completion	'Done' at delivery	Continuous retraining & updates required
Management Style	Plan-driven	Adaptive, experiment-driven

7. Results and Discussion

7.1. The AI Scaling Navigator Framework

AI Scaling Navigator is a six-step, extensively detailed method that has been constructed to handle the multi-faceted challenges that routinely deter piloting efforts involving AI from going viral across the company. Constructed from new research and simultaneous practitioner feedback, the step-by-step method of the Navigator facilitates incremental construction while still allowing room for iterative tuning over the course of time as the environment changes:

- Pilot Discovery: Identify high-impact and strategically important use cases. Short-list candidate use cases based on ROI, feasibility on current infrastructure, and ability to address high-priority business requirements.
- Data & Talent Readiness: Assess data sourcing, data quality, availability, and governance. Set up cross-functional teams that combine data science, engineering, and domain knowledge in order to enable easy execution.
- Executive Sponsorship: Secure public leadership sponsorship from executive leaders themselves and formal budget support. Repeat executive attention underlies resource delegations and prioritization.
- Operationalization (MLOps): Construct automated pipelines from development through to testing, deployment, and monitoring, and then returning to reduce risk and maximize consistency.
- Business Alignment & Change Management: Align AI outcomes into business operations, train, and manage cultural resistance well before.
- Scalable Deployment & Optimization: Scale up winning pilots throughout the entire organization, monitor performance, and optimize on a regular basis.

7.2. AI Pilot Failure Rates and Root Causes

Table 2: AI Pilot Failure Rates and Main Causes

Source	Year	Failure Rate	Main Causes
IDC	2025	88%	Talent shortages, insufficient data readiness.
Gartner	2025	85%	Poor data quality, weak business alignment.
McKinsey	2023	70%	Integration challenges, lack of sustained sponsorship.
S&P Global	2024	~80%	Resource allocation issues, leadership gaps.

It can be seen from the benchmark information that technological restrictions are only part of the problem; organizational readiness, leadership endorsement, and resource focus are equally crucial [1-4].

7.3. Sector Case Vignettes: Scaling in Practice

- Retail Sector Global retail chain piloted an AI-based demand forecasting platform with enhanced accuracy to ±15%. Hub consolidation, training, and executive sponsorship facilitated rollout to 25 hubs, decreasing out-of-stocks by 30% and waste by 18% [12].
- Manufacturing Sector Predictive maintenance pilot rollouts decreased downtime by 25%. Centralized MLOps and change management piloted up to 15 factories, reduced downtime by half, and increased equipment lifespan by 12% [5].
- Financial Services Credit risk analysis using AI improved default forecasting. Early bias and compliance detection facilitated scaling to all loan product types, enhanced regulatory scores, and decreased processing time by 20% [6].

7.4. Research Gaps and Practitioner Needs

Table 3: Research Gaps and Practitioner Needs

Research Gap	Practitioner Need
Model retraining in dynamic environments	Automated drift detection and scheduled retraining triggers.
Linking technical success to business KPIs	Unified KPI dashboards capturing operational and financial metrics.
Measuring organizational AI maturity	Standardized benchmarks to assess readiness and progress

Closing these gaps requires tighter integration across engineering, business, and compliance teams, and strong tooling to keep abreast over time against equity, maturity, and business impact.

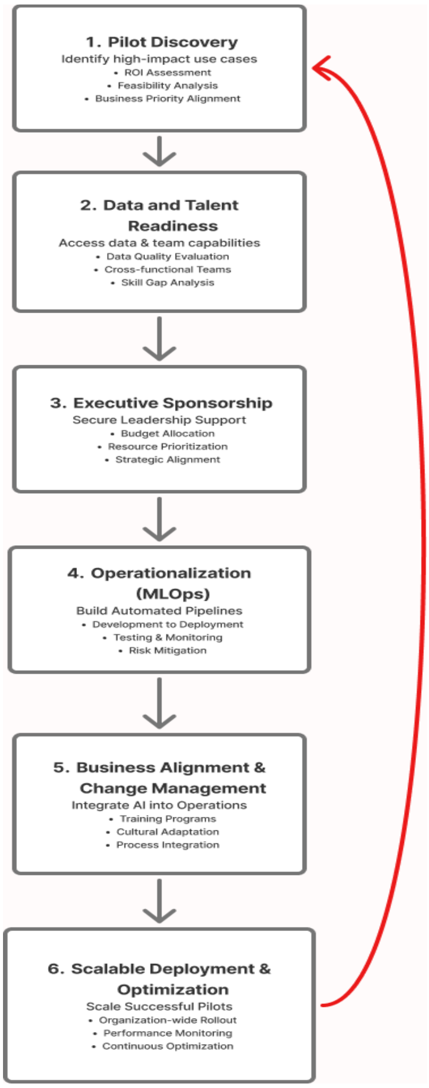

Fig 1: The AI Scaling Navigator Six Phases of Adoption

Flowchart illustrating the six sequential steps to scale AI from pilots to enterprise-wide usage.

8. Future Research

The AI boom in the business world requires ongoing studies to address current gaps and predict upcoming issues. Based on current trends and recent research, several emergent key areas require more scholarly contributions and pragmatic relevance:

8.1. Organizational AI Maturity Models

Even though initial ‘AI-readiness’ maturity frameworks have been suggested, these are sometimes not cross-industry or cross-firm-size validated. More research must develop and empirically test exhaustive maturity models for cultural, technical, governance, and ethics. Long-term, longitudinal studies must be conducted in order to relate organization maturity with scaling pilots to enterprise-scale deployments.

8.2. Measuring Value and Technical Metrics

One of the chronic gaps in practice and research is quantitative correlation of model performance with business effect. Technical metrics (e.g., F1-score, AUC) or anecdotal financial outcomes are provided by most studies, but do not compete with systematic ROI estimation models. The future may adopt cost-benefit analysis tools, together with causal inference procedures, to bridge technical improvements into successfully compelling financial or operational KPIs.

8.3. AI Governance, Regulation, and Ethics

As global attention is on ethical AI, subsequent research must have integrations of explainability, fairness, and ethical governance as first-class project outcomes instead of second-class add-ons. Subsequent regulations (e.g., EU AI Act) will need robust enforcement mechanisms. Cross-industry comparison of how firms apply ethical 'gates' throughout the AI life cycle would provide useful, actionable advice.

8.4. Talent & Human Capital Development

The need for 'translator' or bilingual AI-jobs is pressing, particularly for scalable projects. Future research can compare the effectiveness of internal training academies, rotational programs, and university-industry programs in developing the new, cross-functional skills required by AI transformation. Effective talent models those with business, data, and product expertise should be prioritized via case studies.

8.5. Maintaining AI with Abundant MLOps

Most of the firms are in the initial stages of implementing MLOps. Studies need to explore scalable and automated methods for retraining models and monitoring performance in complex real-world environments, especially the implementation of explainable AI (XAI) [8], continuous validation pipelines, and AI-safe layers for preventing model drift and bias.

8.6. Sharing Benchmarks and Case Studies

There is a lack of visibility into overall project timelines, issues, and cost-to-scale. Projects allowing anonymous, cross-industry benchmarking collating information on cycle times, typical roadblocks, and best practices would enable organizations to develop realistic expectations and roadmaps.

8.7. Emerging Technologies and Methods

Subsequent research needs to investigate the impact of emerging trends, e.g., edge AI deployment, federated learning, and AutoML on scaling drivers, organizational design, and governance requirements. Early indications are that these technologies deepen as well as alleviate some of the scale-up challenges today.

9. Conclusion

Scaling pilots to enterprise-wide solutions necessitates something bigger than tech success. There needs to be strong leadership agreement, aligned strategy, and ops practices embedded. The AI Scaling Navigator presented is a disciplined, repeatable framework tackling industry-shared bottlenecks.

- Practice Implications the use of systematic methods, investment in cross-functional competencies, and sophisticated business alignment can reduce pilot attrition rates by a significant amount.
- Limitations the results mainly rely on medium- to-large companies; data on small company contexts are unavailable.
- Future Research Quantitative validation of Navigator phase phenomena, comprehensive studies, and additional scale studies on controlled industries.

9.1. Conflicts of Interest

The authors declare that they have no conflicts of interest or competing interests in relation to the research, analysis, and publication of this article. All analysis and conclusions were generated independent of employers and in reference to organizations; the views expressed are the authors and should not be understood to necessarily represent the views of the University of Southern California, or any mentioned organization. No proprietary or customer data were used; case vignettes were anonymized, and personally identifiable information was neither collected nor disclosed.

9.2. Acknowledgements

The authors would like to express gratitude to companies such as IDC, Gartner, and Boston Consulting Group for their survey statistics that were referenced for this paper. We are also appreciative of the other associations and publications listed in the references, like McKinsey & Company, S&P Global, RAND Corporation, Deloitte, MIT Sloan Management Review, CIO, Forbes, and CFO Dive, for research reports and articles that shaped our understanding of adoption patterns, operational concerns, and success drivers across industries. The authors recognize the pioneering efforts of the authors and practitioners mentioned, whose works informed the analysis presented herein.

References

[1] IDC, AI adoption and maturity survey. 2025. [Online]. Available: https://www.idc-a.org/insights/0bKr4NJQdK5sYcAQaGZD
[2] Gartner, Lack of AI-ready data puts AI projects at risk. 2025. [Online]. Available: https://www.gartner.com/en/newsroom/press-releases/2025-02-26-lack-of-ai-ready-data-puts-ai-projects-at-risk
[3] McKinsey & Company, The state of AI in 2023. 2023. [Online]. Available: https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-in-2023-generative-ais-breakout-year
[4] S&P Global, Why most AI projects stall. 2024. [Online]. Available: https://press.spglobal.com/2024-11-19-S-P-Global-Market-Intelligences-New-Outlook-Report-Finds-Enterprise-Adoption-of-Generative-Artificial-Intelligence-at-an-Inflection-Point
[5] Deloitte, AI in manufacturing. 2024. [Online]. Available: https://manufacturingdigital.com/smart-manufacturing/deloitte-smart-factory-transformations-in-manufacturing
[6] MIT Sloan Management Review, AI maturity: From pilots to transformation. 2024. [Online]. Available: https://mitsloan.mit.edu/ideas-made-to-matter/whats-your-companys-ai-maturity-level
[7] Joshi, T., & Bhat, M. (2024). Challenges with Developing and Deploying AI Models and Applications

in Industrial Domains. Progress in T. Joshi and M. Bhat, "Challenges with Developing and Deploying AI Models and Applications in Industrial Domains," Progress in Artificial Intelligence, vol. 2024, 2024. [Online]. Available: https://doi.org/10.1007/s44163-024-00151-2

[8] Kumar and S. Ramesh, "MLOps Challenges in Industry 4.0," SN Computer Science, vol. 2024, 2024. [Online]. Available: https://www.scitepress.org/Papers/2022/115896/115896.pdf

[9] R. Shaikh, F. Ahmed, and M. Niazi, "Artificial Intelligence in Project Management: Opportunities and Challenges," Project Management Journal, vol. 2024, 2024. [Online]. Available: https://www.scirp.org/journal/paperinformation?paperid=136694

[10] E. Schuman, “88% of AI pilots fail to reach production but that's not all on IT,” CIO, Mar. 25, 2025.

[11] Tyson, “Only 4% of companies reap full value from AI: BCG,” CFO Dive, Oct. 24, 2024.

[12] V. Lukic, S. Hess, M. Mocker, S. El Khaouli, and J. Kanig, “Scaling AI Pays Off, No Matter the Investment,” Boston Consulting Group, Jan. 10, 2023. [Online]. Available: https://www.bcg.com/publications/2023/scaling-ai-pays-off

[13] RAND Corporation, “The Root Causes of Failure for Artificial Intelligence Projects and How They Can Succeed,” RR-A2680-1, 2024. [Online]. Available: https://www.rand.org/pubs/research_reports/RRA2680-1.html

[14] P. Bendor-Samuel, “Reasons Why Generative AI Pilots Fail to Move Into Production,” Forbes, Jul. 10, 2023. [Online]. Available: https://www.forbes.com/sites/peterbendorsamuel/2023/07/10/reasons-why-generative-ai-pilots-fail/

[15] R. Edjlali, “Lack of AI-Ready Data Puts AI Projects at Risk,” Gartner, Feb. 26, 2025. [Online]. Available: https://www.gartner.com/en/newsroom/press-releases/2025-02-26-lack-of-ai-ready-data-puts-ai-projects-at-risk

[16] Gupta, S. Kamath, M. Ribeiro, et al., “Longitudinal analysis of ML model degradation in production,” unpublished manuscript, 2023.

[17] Khanfar, S. D. Ali, M. J. Yazdi, and T. Al-Emran, “Factors influencing the adoption of AI systems: A systematic literature review,” Management Decision, 2025.

[18] Masud, M. M., Moniruzzaman, M., Rahman, M. M., & Noor, S. (2009). Effect of poultry manure in combination with chemical fertilizers on the yield and nutrient uptake by chilli in the hilly region. *J. Soil Nat*, *3*(2), 24-27.

[19] Lakshmikanthan, G., & Sreekandan Nair, S. (2024). Protecting Self-Driving Vehicles from attack threats. International Journal of Emerging Research in Engineering and Technology, 5(1), 16-20. https://doi.org/10.63282/y6vhvm05