Sarcouncil Journal of Engineering and Computer Sciences

ISSN(Online): 2945-3585

Volume- 04| Issue- 07| 2025

Research Article **Received:** 10-06-2025| **Accepted:** 05-07-2025 | **Published:** 24-07-2025

Why AI Fails Without Strong Data Engineering and What You Can Do About It

Srikanth Maru
Independent Researcher, USA

Abstract: The artificial intelligence sphere faces a paradoxical reality where unknown investment and technological advancement attend with patient perpetration challenges. This dissociation stems primarily from abecedarian scarcities in underpinning data structure rather than limitations in algorithmic approaches. Organizations constantly underrate the complexity of data engineering conditions for product-grade AI systems, creating cycles of promising aviators that fail to gauge. The document examines how data engineering breakdowns manifest across AI executions, revealing common failure patterns including channel complexity, specialized debt accumulation, and integration challenges. Through artificial case exemplifications gauging manufacturing, consumer electronics, and independent systems, the document illuminates critical architectural principles for erecting AI-ready data platforms. A perpetration roadmap outlines practical approaches for transubstantiating heritage systems, developing organizational capabilities, and measuring success through specialized performance pointers aligned with AI issues.

Keywords: Data Infrastructure, Technical Debt, Model Degradation, Pipeline Architecture, Implementation Roadmap.

INTRODUCTION

The artificial intelligence geography presents a paradoxical reality despite unknown situations of investment and technological advancement, associations continue to struggle with rephrasing AI intentions into practical business value. According to recent industry exploration, high-performing associations are more than twice as likely to report significant value from AI compared to other companies, pressing a growing perpetration gap across sectors (Singla, A. *et al*., 2025). This difference stems not primarily from limitations in model architecture or computational coffers, but rather from abecedarian failings in the underpinning data structure supporting AI enterprise. Organizations constantly underrate the complexity of data engineering conditions for product-grade AI systems, leading to a cycle of promising aviators that fail to gauge.

Data negotiating scarcities manifest throughout the AI development lifecycle, creating slinging failures that undermine model performance and business impact. When examining AI perpetration challenges, the lack of suitable data architectures constantly emerges as a critical hurdle to success. Exploration indicates that associations with mature data engineering practices are mainly more likely to successfully implement AI at scale, while those with fractured data infrastructures struggle to progress beyond experimental stages (Singla, A. *et al*., 2025). The inability to efficiently pierce, clean, and integrate data across organizational boundaries creates backups that no quantum of algorithmic complication can overcome, resulting in extended development cycles and lowered returns on AI investments.

A comprehensive examination of enterprise AI executions reveals intimidating statistics regarding design issues. Organizations report that a significant portion of AI enterprise stalls or fails due to problems with underpinning data channels and structure limitations (Jakubik, J. *et al*., 2024). These failures generally stem from inconsistent data governance practices, shy data quality operation systems, and inadequate integration mechanisms across organizational silos. The Business Information Systems Engineering journal notes that data preprocessing conditioning, including cleaning, transformation, and confirmation, frequently consumes a disproportionate quantum of design coffers, creating substantial delays in model deployment and adding overall perpetration costs (Jakubik, J. *et al*., 2024). These challenges emphasize the critical significance of robust data engineering practices as a prerequisite for AI success.

The abecedarian thesis arising from assiduity experience suggests that successful AI deployment depends on establishing solid data engineering foundations before pursuing advanced modeling methods. This architectural approach prioritizes data quality, availability, and governance as primary determinants of AI issues rather than focusing simply on algorithm selection or computational scaling. The substantiation indicates that associations achieving significant AI-driven business value constantly invest in abecedarian data capabilities before spanning their modeling sweats (Singla, A. *et al*., 2025). This sequence of development, establishing a robust data structure before enforcing sophisticated models, represents a

Copyright © 2022 The Author(s): This work is licensed under a Creative Commons Attribution- NonCommercial-NoDerivatives 4.0 (CC BY-NC-ND 4.0) International License

***Corresponding Author:** Srikanth Maru

DOI- https://doi.org/10.5281/zenodo.16408956

critical success factor constantly overlooked in AI strategy development.

This analysis examines perpetration gestures across three distinct technological disciplines: manufacturing intelligence systems, consumer electronics platforms, and independent seeing technologies. Each sphere presents unique data engineering challenges that illuminate assiduity-agnostic principles for erecting flexible data foundations to support sophisticated AI operations. Manufacturing surroundings face complex challenges related to time-series data integration across distant product systems. Consumer electronics ecosystems struggle withcross-divisional data silos that help holistic client perceptivity. Autonomous systems bear high-perfection detector data estimation and real-time processing capabilities that push the boundaries of conventional data infrastructures (Jakubik, J. *et al*., 2024). Inclusively, these disciplines give a comprehensive view of how data engineering practices directly impact AI perpetration issues across diligence.

THE ANATOMY OF DATA ENGINEERING BREAKDOWNS IN AI PROJECTS

The intricate relationship between data engineering structure and artificial intelligence systems establishes a critical reliance chain that unnaturally determines AI success. Machine literacy systems correspond not simply to algorithms but to complex, connected factors where data collection mechanisms, point birth channels, and serving structure play inversely vital roles as the models themselves. According to an exploration published in Neural Information Processing Systems, these" trap" issues produce scripts where changes to one data channel element can have unanticipated consequences throughout the entire AI system (Sculley, D. *et al*., 2015). This interdependence means that indeed state- of- the- art algorithms can not overcome abecedarian excrescencies in how data is collected, reused, and delivered. Organizations constantly discover this reality too late in development cycles, after substantial coffers have been committed to model development atop unstable data foundations.

Across diligence, recreating failure patterns crop up when examining an unprofitable AI enterprise. The conception of" retired specialized debt" in machine literacy systems reveals several orders of data engineering breakdowns that constantly undermine AI systems' unstable data dependences, channel jungles with complex metamorphoses, point birth challenges, and configuration debt related to how data systems are organized (Sculley, D. *et al*., 2015). These patterns manifest through inconsistent model geste, unanticipated performance decline, and conservation challenges that grow exponentially as systems scale. The exploration identifies" correction falls" where compensatory sense for data quality issues gradually accumulates throughout the channel, creating brittle systems that ultimately collapse under their own complexity. Rather than addressing root causes in data engineering practices, associations constantly apply short-term workarounds that complicate long-term conservation challenges.

Heritage data systems accumulate specialized debt through colorful mechanisms that directly impact AI capabilities. The conception of" noway-ending literacy" described in Dispatches of the ACM highlights how traditional data infrastructures unnaturally discord with ultramodern AI conditions for nonstop knowledge accession and adaptation (Mitchell, T. *et al*., 2018). Traditional systems designed for transactional processing struggle with the exploratory, iterative nature of machine literacy workflows. This architectural mismatch creates scripts where data structure can not evolve alongside AI capabilities, forcing negotiations that limit model complication. The exploration emphasizes how heritage systems generally warrant critical capabilities for tracking data provenance, managing experimental variations, and handling the complex feedback circles essential in ultramodern AI systems.

Poor data architecture opinions waterfall through machine literacy systems with multiplicative negative goods. Research demonstrates how architectural excrescencies in data systems lead to what machine literacy interpreters term" channel debt," where each stage of data processing adds complexity that composites through the system (Sculley, D. *et al*., 2015). This cascading failure pattern creates scripts where putatively minor data engineering opinions made early in development cycles dramatically constrain model capabilities in production. The paper specifically identifies how common practices like duplicated point birth sense, disconnected analysis, and cement law proliferation directly contribute to model performance decline over time. These architectural limitations constantly remain unnoticeable until systems face unanticipated data distributions or scaling conditions.

Copyright © 2022 The Author(s): This work is licensed under a Creative Commons Attribution- NonCommercial-NoDerivatives 4.0 (CC BY-NC-ND 4.0) International License

Publisher: SARC Publisher

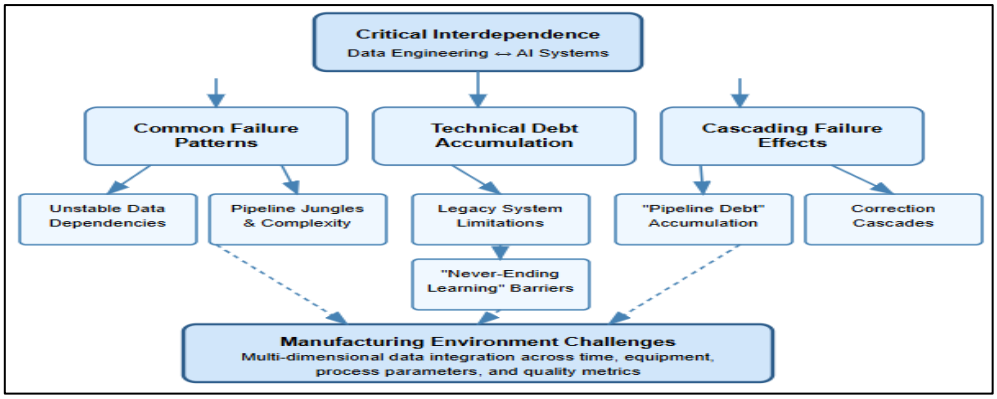


Fig 1: The Anatomy of Data Engineering Breakdowns in AI Projects (Sculley, D. *et al*., 2015; Mitchell, T. *et al*., 2018)

Manufacturing surroundings present particularly grueling exemplifications of dimensionality and integration issues in artificial intelligence operations. The exploration highlights how complex product surroundings induce multi-dimensional data across time, outfit countries, process parameters, and quality criteria that traditional data infrastructures struggle to integrate coherently (Sculley, D. *et al*., 2015). These dimensional challenges produce abecedarian walls to enforcing prophetic conservation, quality optimization, and process control systems. The paper identifies specific failure modes, including underspecified data dependences, correction falls across detector systems, and feedback circles that remain unaccounted for in traditional data channels. Successful manufacturing AI executions bear technical data engineering approaches designed explicitly to address these multi-dimensional integration challenges.

LESSONS FROM THE FIELD

A prominent digital fabrication company's trip toward AI-enabled manufacturing illustrates abecedarian challenges in transubstantiating heritage data channels. This manufacturer originally reckoned on traditional manufacturing execution systems(MES) designed around batch processing paradigms, where product data was collected, reused, and analyzed in diurnal aggregation cycles. Research published in IEEE Deals on Industrial Informatics highlights how similar heritage infrastructures unnaturally discord with the conditions of ultramodern artificial cyber-physical systems, which demand near-nonstop feedback circles between physical infrastructure and decision-making algorithms (Karnouskos, S. *et al*., 2020). The exploration identifies how artificial agents — independent software factors that represent physical realities — bear unnaturally different data infrastructures than traditional manufacturing systems. These cyber-physical systems depend on event-driven communication patterns, decentralized decision-making capabilities, and flexible reconfiguration possibilities that traditional batch-acquainted infrastructures can not support. The metamorphosis trip needed to enforce a multi-agent system architecture where individual product factors were represented as data-generating agents within a coordinated ecosystem. Particularly grueling was the integration of structured product criteria (cycle times, quality measures, material consumption) with unshaped design lines and client feedback. The exploration demonstrates how artificial cyber-physical systems must bridge this structured-unshaped peak through semantic data models that establish meaningful connections between product parameters and consumer-facing labor (Karnouskos, S. *et al*., 2020).

A transnational electronics pot exemplifies the scale and complexity challenges essential in consumer technology ecosystems. With operations gauging mobile bias, home appliances, display technologies, and semiconductor manufacturing, this association faced extraordinary challenges in establishing unified data visibility across product divisions. Research in IEEE Deals on Industrial Informatics examines how agent-grounded infrastructures can address cross-divisional integration challenges through formalized communication protocols and semantic interoperability fabrics (Karnouskos, S. *et al*., 2020). The electronics manufacturer originally operated with independent data operation systems across business units, creating artificial boundaries that averted holistic client experience analysis. The metamorphosis needed to enforce a hierarchical agent-grounded armature where product systems, enterprise operations, and client touchpoints operated as connected but independent agents.

Copyright © 2022 The Author(s): This work is licensed under a Creative Commons Attribution- NonCommercial-NoDerivatives 4.0 (CC BY-NC-ND 4.0) International License

Publisher: SARC Publisher

This approach enabled the association to maintain functional independence while establishing standardized data exchange mechanisms across previously siloed systems. The exploration highlights how artificial cyber-physical systems depend on prostrating heritage integration walls through service-acquainted infrastructures that abstract underlying system complexity. Organizations enforcing similar multi-agent fabrics can stoutly reconfigure data flows without dismembering functional systems, enabling incremental metamorphosis rather than high-threat" big bang" integration approaches (Karnouskos, S. *et al*., 2020).

An independent vehicle sensing technology provider demonstrates the extreme perfection conditions essential for safety-critical perception systems. Research published in arXiv examines the comprehensive challenges in real-time 3D perception for independent driving, pressing how detector data infrastructures must satisfy strict conditions across quiescence, delicacy, and robustness confines contemporaneously (Qian, R. *et al*., 2022). The company's experience developing lidar-grounded perception systems revealed significant limitations in conventional data engineering approaches when applied to high-dimensional spatial data aqueducts. The exploration identifies specific architectural considerations for 3D perception channels, including detector configuration optimization, estimation methodologies, and evaluation criteria that enable effective machine literacy on spatial-temporal data. Particularly grueling was establishing robust ground truth data for model training, taking sophisticated detector emulsion across multiple modalities, including lidar, radar, and camera systems. The exploration examines how perception system infrastructures must accommodate different environmental conditions, including adverse rainfall, varying illumination, and complex multi-object scripts (Qian, R. *et al*., 2022). The company enforced sophisticated estimation channels that continuously covered detector performance across environmental variations, automatically relating drift conditions, and requiring recalibration. The exploration emphasizes how independent perception systems depend on comprehensive confirmation methodologies that totally estimate performance across both common and edge-case scripts, taking technical data infrastructures that can efficiently store, recoup, and process massive multi-sensor datasets.

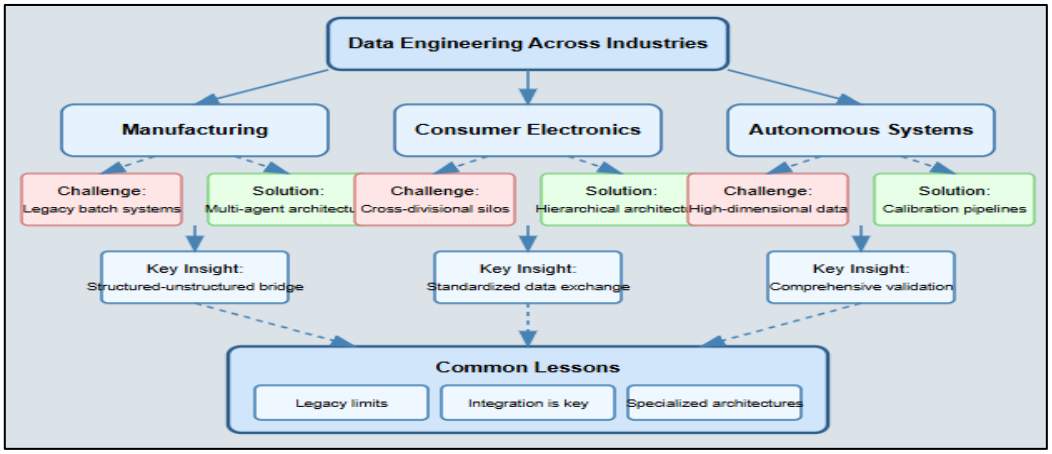


Fig 2: Lessons from the Field (Karnouskos, S. *et al*., 2020; Qian, R. *et al*., 2022)

ARCHITECTURAL PRINCIPLES FOR AI-READY DATA PLATFORMS

Designing scalable data platforms for artificial intelligence operations requires careful consideration of scaling approaches that align with workload characteristics. Research published in arXiv examines abecedarian architectural patterns that impact scalability in deep literacy systems, pressing the distinctions between vertical and perpendicular scaling strategies (Mayer, R. *et al*., 2020). The paper identifies how the essential community in neural network infrastructures creates openings for distributed training across multiple calculation nodes, while contemporaneously introducing challenges related to communication outflow and synchronization conditions. Vertical scaling approaches distribute workloads across homogeneous bumps, enabling direct out-of-the-box advancements for embarrassingly parallel operations such as hyperparameter tuning and inference serving. Vertical spanning strategies, by discrepancy, influence decreasingly important individual bumps

Copyright © 2022 The Author(s): This work is licensed under a Creative Commons Attribution- NonCommercial-NoDerivatives 4.0 (CC BY-NC-ND 4.0) International License

Publisher: SARC Publisher

with enhanced memory and processing capabilities, reducing communication outflow for grade-ferocious operations. The exploration presents a taxonomy of scaling patterns observed across product AI deployments, including model community, data community, and cold-blooded approaches that combine both strategies grounded on neural architecture characteristics. The study further examines how scaling opinions directly impact training confluence rates, with certain distribution patterns introducing stochastic goods that modify optimization gestures compared to single-knot executions (Mayer, R. *et al*., 2020).

Data quality governance fabrics give essential rails for machine literacy systems, with structured approaches to quality operation directly impacting model performance and trustworthiness. Research from arXiv presents architectural patterns that totally address data quality enterprises throughout AI development lifecycles (Mayer, R. *et al*., 2020). The paper identifies specific quality constraints particularly applicable to machine learning workloads, including absoluteness, thickness, delicacy, and punctuality — constraints that extend traditional data quality frameworks with ML-specific considerations such as point drift, class imbalance, and boundary condition representation. The study introduces the concept of" data confirmation contracts" that formalize quality prospects throughout the channel, enabling automated discovery of distribution shifts that might impact model performance. Particularly notable is the examination of how quality considerations differ across development stages, with early experimental phases prioritizing diversity and representation while product deployments emphasize thickness and stability. The exploration details architectural approaches for enforcing nonstop data confirmation, including distributed statistical analysis of incoming data aqueducts, automated schema enforcement, and technical anomaly discovery for multivariate point spaces. The paper further explores how data quality fabrics must evolve beyond point-in-time confirmation to nonstop monitoring across complete data lifecycles (Mayer, R. *et al*., 2020).

Metadata operation represents a foundational capability for resolvable and governable artificial intelligence systems. Research published in the Asian Journal of Information Science and Technology examines how comprehensive metadata shadowing enables critical AI capabilities, including lineage attestation, model explainability, and compliance verification (Oyighan, D. *et al*., 2024). The paper explores how metadata conditions for AI systems extend significantly beyond traditional data registers, challenging shadowing of experimental configurations, model infrastructures, hyperparameter selections, and training data characteristics. This expanded metadata compass enables traceability from specific prognostications back to source data, metamorphosis sense, and model training opinions capabilities increasingly needed by nonsupervisory fabrics governing automated decision systems. The study presents architectural patterns for enforcing metadata depositories that accommodate the different vestiges generated throughout machine literacy lifecycles, including structured roster approaches, graph-grounded relationship models, and allied metadata infrastructures that gauge organizational boundaries. Particularly notable is the examination of how metadata standardization facilitates cross-team collaboration while reducing duplication of point engineering sweats. The exploration further details how comprehensive metadata operation directly enables model governance capabilities, including blessing workflows, interpretation control, and deployment instrument (Oyighan, D. *et al*., 2024).

Ultramodern AI architectures increasingly bear sophisticated integration between batch and streaming processing paradigms to support different logical patterns. Research from arXiv examines architectural approaches for enforcing mongrel processing capabilities that combine batch processing's comprehensive analysis with streaming's real-time responsiveness (Mayer, R. *et al*., 2020). The paper details how machine literacy workloads generally gauge both paradigms, with literal batch processing essential for comprehensive model training while streaming capabilities enable real-time conclusion, nonstop confirmation, and online literacy. The study presents reference infrastructures for enforcing unified processing fabrics that maintain thickness across batch and streaming paths, examining approaches including lambda armature, kappa armature, and specialized variants optimized for machine learning workloads. The exploration identifies specific architectural challenges in maintaining thickness between batch and streaming results, including schema elaboration operation, point calculation alignment, and conciliation of temporal boundaries. The paper further explores how ultramodern sluice processing fabrics have evolved to support

Copyright © 2022 The Author(s): This work is licensed under a Creative Commons Attribution- NonCommercial-NoDerivatives 4.0 (CC BY-NC-ND 4.0) International License

Publisher: SARC Publisher

complex stateful operations, preliminarily limited to batch surroundings, enabling architectural simplification through sluice-first approaches supplemented by periodic batch processing for comprehensive retraining (Mayer, R. *et al*., 2020).

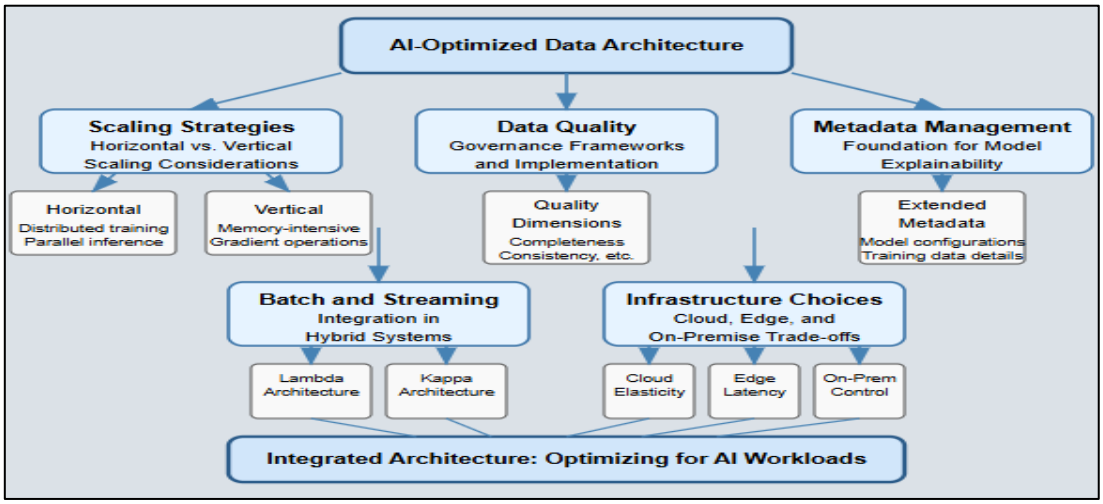


Fig 3: Architectural Principles for Al-Ready Data Platforms (Mayer, R. *et al*., 2020; Oyighan, D. *et al*., 2024)

Structure deployment opinions increasingly shape AI capabilities, with exploration from the Asian Journal of Information Science and Technology examining the evolving relationship between centralized cloud, distributed edge, and traditional on-premise paradigms (Oyighan, D. *et al*., 2024). The paper explores how these deployment models present distinct trade-offs across confines, including data graveness, quiescence conditions, sequestration constraints, and computational viscosity. Edge calculating approaches detect processing capabilities near data generation sources, minimizing quiescence for time-sensitive operations while reducing bandwidth conditions for data-intensive workloads. Pall deployments give pliancy and technical tackle access while polarizing governance and enabling cross-regional model training. On-premises results maintain maximum control over sensitive data while potentially limiting access to technical AI acceleration tackle. The exploration examines cold-blooded deployment patterns that strategically distribute workloads based on specific conditions, including training/ conclusion splits where comprehensive model training occurs in centralized surroundings while conclusion deploys to distributed edge locales. The paper further explores how containerization and unity technologies enable harmonious model deployment across miscellaneous structure surroundings, easing cold-blooded approaches that optimize for both performance and governance conditions (Oyighan, D. *et al*., 2024).

IMPLEMENTATION ROADMAP: FROM LEGACY TO AI-OPTIMIZED

Transforming heritage data systems into AI-ready platforms requires comprehensive assessment methodologies that estimate structural capabilities against unforeseen conditions. Research published in the Journal of Computer Science and Technology presents a methodical approach to assessing big data systems grounded on architectural patterns and functional characteristics (Wu, L. *et al*., 2015). The paper introduces a bracket frame that examines data systems across multiple confines, including processing models(batch vs. sluice), storehouse paradigms(row vs. column-oriented), and query prosecution strategies. This assessment methodology enables associations to evaluate current capabilities against conditions for AI workloads, relating specific gaps rather than pursuing general modernization objectives. The exploration emphasizes how successful metamorphosis enterprises begin with detailed architectural evaluation concentrated on critical capabilities, including distributed processing, resemblant prosecution, fault forbearance, and data position optimization. The assessment phase should totally validate current state limitations, including processing backups, spanning constraints, and integration challenges that would stymie AI perpetration (Wu, L. *et al*., 2015).

Gradational migration approaches significantly reduce the perpetration threat when transitioning from heritage to AI-optimized data platforms. Research presented at the Conference on Machine

Copyright © 2022 The Author(s): This work is licensed under a Creative Commons Attribution- NonCommercial-NoDerivatives 4.0 (CC BY-NC-ND 4.0) International License

Publisher: SARC Publisher

Learning Operations Systems explores migration strategies specifically designed for MLOps metamorphoses (Hanchuk, D. O. 2025). The paper introduces a structured migration frame that organizes metamorphosis conditioning into distinct phases: foundation structure, capability perpetration, integration establishment, and functional transition. This phased approach enables associations to maintain business durability while incrementally introducing new capabilities. The exploration emphasizes how sphere-driven migration strategies that prioritize specific business capabilities rather than specialized factors enable earlier value demonstration and stakeholder alignment. Particularly effective are perpetration patterns that establish concurrence between heritage and new systems during transition ages, allowing gradational workload migration without disruptive cutover events (Hanchuk, D. O. 2025).

Structuring organizational capabilities for AI data engineering requires structured approaches to specialized skill development. Research from the Journal of Computer Science and Technology examines how big data system architecture directly influences needed specialized capabilities (Wu, L. *et al*., 2015). The paper identifies technical knowledge disciplines essential for ultramodern data platforms, including distributed system design, relevant processing optimization, and miscellaneous data integration. This capability mapping enables associations to develop targeted training programs and reclamation strategies aligned with architectural conditions. The exploration emphasizes how successful metamorphosis enterprises establish capability development roadmaps that evolve specialized chops in line with structure perpetration, icing organizational readiness as new systems come available (Wu, L. *et al*., 2015).

Resource allocation strategies significantly impact metamorphosis issues. Research on MLOps maturity models presents resource optimization approaches specifically designed for a machine literacy structure enterprise (Hanchuk, D. O. 2025). The paper introduces a resource planning frame that aligns allocation with specific metamorphosis phases, enabling further effective distribution of technical chops across design lifecycles. This approach recognizes how resource conditions shift as enterprises progress from assessment through perpetration to functional transition. The exploration emphasizes how successful metamorphosis enterprises establish flexible resource models that combine endless core brigades with stoutly allocated specialists grounded on phase-specific conditions (Hanchuk, D. O. 2025).

Measuring data engineering effectiveness requires technical performance pointers aligned with AI issues. Research from both sources emphasizes how traditional IT criteria are deficient in capturing data engineering value in AI surroundings (Wu, L. *et al*., 2015; Hanchuk, D. O. 2025). The studies identify technical pointers, including data integration effectiveness, recycling scalability, and model serving quiescence, that more effectively prognosticate AI perpetration success. Effective dimension fabrics establish clear connections between specialized criteria and business issues, enabling stakeholders to understand how structured investments directly enable AI capabilities. These fabrics apply metric scales that connect functional pointers to strategic objects, easing clearer communication regarding metamorphosis progress (Wu, L. *et al*., 2015; Hanchuk, D. O. 2025).

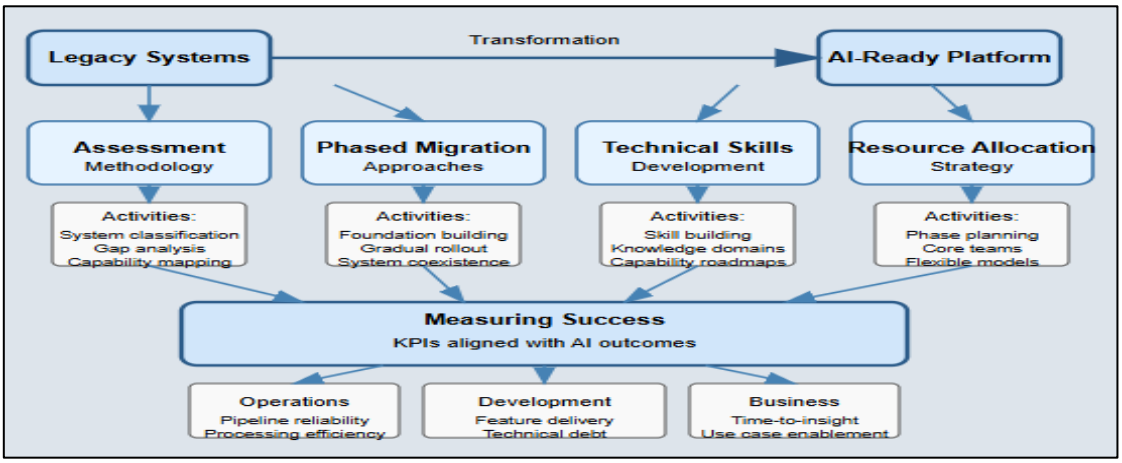


Fig 4: Implementation Roadmap: From Legacy to AI-Optimized (Wu, L. *et al*., 2015; Hanchuk, D. O. 2025)

Copyright © 2022 The Author(s): This work is licensed under a Creative Commons Attribution- NonCommercial-NoDerivatives 4.0 (CC BY-NC-ND 4.0) International License

Publisher: SARC Publisher

CONCLUSION

The trip from heritage data systems to AI-optimized platforms requires deliberate architectural opinions that prioritize data quality, availability, and governance over algorithmic complication. A successful AI enterprise constantly demonstrates that robust data engineering foundations determine issues more definitively than model selection or computational scaling. Organizations must approach metamorphosis through phased migration strategies that balance specialized excellence with business durability, gradually shifting workloads while maintaining functional stability. As data engineering for AI continues evolving, patterns arising from automated quality governance, unified metadata operation, and mongrel processing fabrics will increasingly define perpetration success. Organizations can begin metamorphosis by conducting comprehensive structure assessments, establishing data quality nascences, enforcing metadata governance fabrics, and developing technical criteria that connect specialized capabilities to business issues. The development of data engineering practices specifically optimized for AI workloads represents the critical elaboration necessary to bridge the patient gap between AI eventuality and practical perpetration.

REFERENCES

1. Singla, A., Sukharevsky, A., Yee, L., Chui, M., & Hall, B. "The state of AI: How organizations are rewiring to capture value." *McKinsey & Company* 12 (2025).
2. Jakubik, J., Vössing, M., Kühl, N., Walk, J., & Satzger, G. "Data-centric artificial intelligence." *Business & Information Systems Engineering* 66.4 (2024): 507-515.
3. Sculley, D., Holt, G., Golovin, D., Davydov, E., Phillips, T., Ebner, D., & Dennison, D. "Hidden technical debt in machine learning systems." *Advances in neural information processing systems*. 28 (2015).
4. Mitchell, T., Cohen, W., Hruschka, E., Talukdar, P., Yang, B., Betteridge, J., & Welling, J. "Never-ending learning." *Communications of the ACM* 61.5 (2018): 103-115.
5. Karnouskos, S., Leitao, P., Ribeiro, L., & Colombo, A. W. "Industrial agents as a key enabler for realizing industrial cyber-physical systems: Multiagent systems entering industry 4.0." *IEEE Industrial Electronics Magazine* 14.3 (2020): 18-32.
6. Qian, R., Lai, X., & Li, X. "3D object detection for autonomous driving: A survey." *Pattern Recognition* 130 (2022): 108796.
7. Mayer, R., & Jacobsen, H. A. "Scalable deep learning on distributed infrastructures: Challenges, techniques, and tools." *ACM Computing Surveys (CSUR)* 53.1 (2020): 1-37.
8. Oyighan, D., Ukubeyinje, E. S., David-West, B. T., & Oladokun, B. D. "The role of AI in transforming metadata management: Insights on challenges, opportunities, and emerging trends." *Asian Journal of Information Science and Technology* 14.2 (2024): 20-26.
9. Wu, L., Yuan, L., & You, J. "Survey of large-scale data management systems for big data applications." *Journal of computer science and technology* 30.1 (2015): 163-183.
10. Hanchuk, D. O., & Semerikov, S. O. "Implementing MLOps practices for effective machine learning model deployment: A meta synthesis." *CEUR Workshop Proceedings*. (2025)

Source of support: Nil; **Conflict of interest:** Nil.

Cite this article as:
Maru, S. "Why AI Fails Without Strong Data Engineering and What You Can Do About It" *Sarcouncil Journal of Engineering and Computer Sciences* 4.7 (2025): pp 1132-1139.

Copyright © 2022 The Author(s): This work is licensed under a Creative Commons Attribution- NonCommercial-NoDerivatives 4.0 (CC BY-NC-ND 4.0) International License

Publisher: SARC Publisher