Journal of Management and Economics

TYPE Original Research
PAGE NO. 30-35

Check for updates

From Business Intelligence to Intelligent Automation: Integrating Adaptive Data Engineering and Large Multimodal Models for Competitive Advantage

OPEN ACCESS

SUBMITED 01 September2025
ACCEPTED 16 September2025
PUBLISHED 30 September 2025
VOLUME Vol.05 Issue 06 2025

Dr. Talvion A. Strelnikov
Independent Researcher, Business Analytics, Intelligent Automation & Enterprise Transformation, Novosibirsk, Russia

Abstract: Background: The landscape of Business Intelligence (BI) is undergoing a paradigm shift, moving from historical data visualization to predictive, generative insights powered by Artificial Intelligence. While large enterprises leverage these tools for competitive advantage, the integration of Large Language Models (LLMs) and Large Multimodal Models (LMMs) into standard data pipelines remains a complex engineering challenge, particularly regarding Extract, Transform, Load (ETL) automation.

Methods: This study employs a comprehensive synthesis of recent literature (2020-2025) and theoretical modeling using the TOE-DOI framework. We analyze the efficacy of Adaptive Data Engineering (LLM-ADE) and compare intelligent automation tools (Dataverse vs. TPOT) in the context of optimizing BI workflows.

Results: The analysis indicates that traditional BI success factors are being superseded by the capability to automate unstructured data processing. The integration of LLM-ADE reduces data preparation latency by significant margins, enabling real-time competitive bidding advantages. Furthermore, the adoption of LMMs facilitates a deeper understanding of complex, multimodal market signals that traditional algorithms miss.

Conclusion: We conclude that the future of competitive advantage lies not merely in data access, but in the intelligent automation of the data lifecycle. Organizations that transition to adaptive, LLM-driven architectures will dominate proprietary information landscapes, though

they must navigate new risks related to algorithmic hallucinations and cost management.

Keywords: Business Intelligence, Adaptive Data Engineering, Large Language Models, Intelligent Automation, Competitive Advantage, ETL Optimization, Multimodal AI

1.INTRODUCTION

The accumulation of data assets has long been recognized as a cornerstone of strategic management. Since the mid-1990s, scholars have argued that the development of robust IT assets is essential for long-term competitiveness [4]. However, the definition of these assets has evolved dramatically. In the early 2000s, Business Intelligence (BI) was primarily concerned with the management of knowledge within virtual enterprises, focusing on structured databases and human-centric query logic [1]. Today, the proliferation of unstructured data and the advent of Generative AI have fundamentally altered this landscape. The contemporary enterprise does not suffer from a lack of information, but rather from an inability to synthesize it rapidly enough to inform decision-making in real-time.

This creates a phenomenon known as the "Data Engineering Bottleneck." Traditional Extract, Transform, Load (ETL) pipelines are rigid, rule-based systems designed for structured data. They are ill-equipped to handle the nuance and variability of the data types that drive modern insights—such as natural language text, images, and sensor data. As organizations attempt to leverage Large Language Models (LLMs) to gain a competitive edge, they frequently encounter the limitation that their data infrastructure cannot feed these models effectively [12].

The emergence of Large Multimodal Models (LMMs) represents a "New Era in Human Factors Engineering" [9], offering the potential to process and reason across different modalities. Yet, the adoption of these advanced tools is uneven. While tech giants leverage BI for massive competitive advantages [10], Small and Medium Enterprises (SMEs) struggle with the foundational adoption of Data Management as a Service (DMaaS) and Big Data as a Service (BDaaS) due to cost and technical complexity [5, 7].

This article aims to bridge the gap between traditional BI frameworks and the emerging discipline of Adaptive Data Engineering. By analyzing the intersection of intelligent automation in ETL processes [13] and the generative capabilities of LLMs [11], we explore how organizations can transition from static reporting to dynamic, AI-driven intelligence systems. We examine the hypothesis that the integration of LLM-ADE (Large Language Models with Adaptive Data Engineering) provides a distinct advantage in competitive bidding and proprietary information management [3], fundamentally reshaping the economics of decision support.

2. LITERATURE REVIEW

2.1 The Evolution of Business Intelligence

The trajectory of BI has historically been defined by the capability to structure chaos. Early research by Jermol et al. [1] emphasized the human component of "lessons learned" in virtual enterprises, suggesting that technology was merely a facilitator for human wisdom. This view was supported by Oyku et al. [2], who identified that BI success was contingent not just on software capabilities, but on the decision environment itself. In this "BI 1.0" era, success was measured by the accuracy of historical reporting.

2.2 The Challenge of Adoption in SMEs

As BI technologies matured into cloud-based services, the divide between large enterprises and SMEs widened. Zide and Jokonya [5] highlighted that factors affecting the adoption of Data Management as a Service (DMaaS) in SMEs were deeply rooted in organizational readiness and perceived cost-benefits. Similarly, Wessels and Jokonya [7] found that while Big Data offered theoretical advantages, the practical implementation in smaller firms was hindered by a lack of technical expertise. This suggests that for advanced BI to become ubiquitous, the barrier to entry—specifically the complexity of data engineering—must be lowered through automation.

2.3 The Generative Shift and Intelligent Automation

The introduction of LLMs has disrupted the linear progression of BI tools. Ege et al. [11] explored "ChatGPT as an inventor," revealing that while current LLMs exhibit weaknesses in physical engineering design, they possess remarkable capabilities in conceptual synthesis. This generative capability extends to data processing itself. Mantri [13] provided a comparative study of intelligent automation in ETL processes, specifically examining tools like Dataverse and TPOT. The findings suggest that we are moving toward a "self-healing" data pipeline, where AI systems can identify and rectify data quality issues without human intervention.

2.4 Large Multimodal Models (LMMs)

Fan et al. [9] argue that we are entering an era of LMMs, where systems can process visual, auditory, and textual data simultaneously. This capability is critical for modern competitive intelligence. In competitive bidding environments [3], proprietary information is often hidden in non-structured formats—satellite imagery of supply chains, video analysis of competitor products, or sentiment analysis of executive speeches. Traditional BI tools are blind to these signals; LMMs are not.

3. METHODOLOGY

To analyze the transformation of BI systems, this article utilizes the Integrated TOE-DOI Framework (Technology-Organization-Environment / Diffusion of Innovation), as applied in recent public sector reform studies by Basloom et al. [6]. The TOE framework is particularly appropriate for analyzing complex technological adoption because it accounts for internal readiness and external pressures.

- Technology Context: We analyze the shift from rigid ETL to Adaptive Data Engineering (ADE) and the specific capabilities of LMMs [9].
- Organization Context: We examine the internal capabilities required to support these systems, drawing on Ross et al.'s [4] concept of IT assets and the changing role of human analysts [11].
- Environment Context: We consider the competitive pressure described by Richard et al. [3] regarding proprietary information and bidding strategies.

Our analysis synthesizes technical specifications and case study data from recent literature (2023-2025), specifically focusing on the architectural models proposed by Choi and Gazeley [12] and the case studies of tech giants provided by Patel [10]. This synthesis allows for a theoretical reconstruction of an optimal "Modern BI Stack" that integrates these disparate technologies.

4. RESULTS:

The Architecture of Competitive Advantage

The synthesis of the collected literature points toward a distinct architectural evolution. The traditional data warehouse, fed by manual ETL scripts, is being replaced by an ecosystem characterized by Adaptive Data Engineering (ADE). This section details the components of this new architecture and its performance implications.

4.1 Adaptive Data Engineering (LLM-ADE)

The concept of LLM-ADE, as introduced by Choi and Gazeley [12], represents the most significant finding in our review of recent technical advancements. Traditional data engineering requires predefined schemas; if the data format changes, the pipeline breaks. LLM-ADE utilizes the reasoning capabilities of Large Language Models to interpret incoming data dynamically.

When an LLM-ADE system encounters a new data source—for example, a competitor's unstructured PDF financial report—it does not reject the file for lacking a schema. Instead, the LLM analyzes the document structure, infers the relevant schema, generates the extraction code on the fly, and normalizes the data into the analytical warehouse. This "schema-on-read" capability, powered by generative AI, reduces the time-to-insight from days to minutes.

This capability directly addresses the adoption barriers cited by Zide and Jokonya [5]. SMEs often lack the dedicated data engineering teams to maintain brittle ETL pipelines. An adaptive system, which requires less manual maintenance, effectively lowers the "tax" of BI adoption, democratizing access to high-level analytics.

4.2 Intelligent Automation of ETL: Dataverse vs. TPOT

Mantri's comparative study [13] provides crucial data regarding the implementation of these pipelines. The comparison between Dataverse (a managed data platform) and TPOT (Tree-based Pipeline Optimization Tool) reveals a dichotomy in the current market.

- TPOT (Automated Machine Learning): The study indicates that tools like TPOT excel in optimizing the transformation logic itself. By using genetic programming, TPOT can automatically select the best feature engineering steps and model parameters. In the context of BI, this means the system can "learn" which data features are most predictive of business success (e.g., sales churn, inventory depletion) without a human analyst explicitly defining those relationships.
- Dataverse: Conversely, platforms like Dataverse offer stronger governance and integration, particularly for organizations already embedded in the Microsoft ecosystem. However, the study suggests that for LLM deployment—where the goal is to feed a generative model—the flexibility of Python-based automated tools (like TPOT or custom LLM-ADE scripts) often outperforms rigid enterprise platforms in terms of adaptability to novel data types.

4.3 The Role of LMMs in Environmental Scanning

Fan et al. [9] describe the capabilities of Large Multimodal Models (LMMs). In a competitive bidding scenario [3], the ability to analyze multimodal data is a critical differentiator. Consider a construction firm bidding on a government contract.

- Traditional BI: Analyzes past bid amounts and public budget documents (Structured Text/Numbers).
- LMM-Enhanced BI: Can ingest the Request for Proposal (RFP) documents (Text), analyze site survey drone footage (Video/Image) to estimate terrain challenges, and listen to municipal council meeting recordings (Audio) to detect political sentiment regarding the project.

The integration of these modalities creates a "Rich Feature" set that allows the bidding firm to price their offer more accurately, minimizing the "Winner's Curse" (overpaying to win a bid) often discussed in economic literature [3]. This represents a tangible ROI from the adoption of LMMs.

4.4 Cost-Benefit Considerations

Härting and Sprengel [8] emphasize the necessity of rigorous cost-benefit analysis for data analytics. The move to LLM-driven BI is capital intensive. Inference costs for high-performance models (like GPT-4 or Claude 3 Opus) are non-trivial.

Our analysis suggests a U-shaped cost curve.

1. Initial Phase: High costs due to infrastructure setup and model fine-tuning.

2. Operational Phase: Costs decrease as LLM-ADE automates manual engineering hours. The reduction in human labor (data cleaning, schema mapping) offsets the compute costs.

3. Scaling Phase: Costs may rise again as the volume of processed unstructured data explodes.

Therefore, organizations must implement "Model Routing"—using cheaper, smaller models for routine ETL tasks and reserving powerful LMMs for complex strategic inference.

4.5 Deep Dive: The Competitive Bidding Algorithm

Integrating the insights from Richard et al. [3] regarding proprietary information with the technical capabilities described by Patel [10], we can construct a theoretical model for AI-driven competitive bidding. In traditional economics, proprietary information is static. In an AI-driven model, proprietary information is synthesized.

The algorithm functions as follows:

1. Ingest: The system ingests public market data.

2. Enrich: LLM-ADE enriches this with "Grey Data" (semi-public data like employee reviews, patent filings, Github commits).

3. Simulate: The system uses a Generative Agent approach (similar to Ege et al. [11] in invention) to simulate competitor responses. "If we bid X, how likely is Competitor Y to counter-bid based on their historical aggression and current financial health?"

4. Optimize: The system recommends a bid price that maximizes the probability of winning while preserving margin.

This transition from descriptive analytics ("What did we bid last time?") to prescriptive simulation ("What should we bid to beat Competitor Y?") is the defining characteristic of the new era of Business Intelligence.

4.6 The Human Element in the Loop

Despite the automation described above, the human factor remains pivotal. As noted in the early work by Jermol [1], lessons learned are a human construct. Ege et al. [11] highlight that while LLMs are excellent at combinatorial creativity, they lack "grounding" in physical reality. In a BI context, an LLM might spot a correlation that makes statistical sense but is operationally impossible.

Therefore, the role of the BI professional shifts from "Report Writer" to "Model Auditor." The skillset required moves away from writing SQL queries and toward understanding model interpretability, bias detection, and prompt engineering. This aligns with the "New Era in Human Factors" described by Fan [9], where the interface between human intent and machine execution becomes the primary bottleneck.

4.7 Institutional Isomorphism and the Tech Giants

Patel's analysis [10] of tech giants reveals a phenomenon of institutional isomorphism—as Google, Amazon, and Microsoft adopt these AI-driven BI practices, they set a standard that forces the rest of the market to adapt. This creates significant pressure on SMEs. The "factors affecting adoption" [5] are no longer just internal preferences but external survival mandates. If a small logistics company competes against Amazon's logistics network, which is optimized by LMMs and real-time adaptive engineering, the SME is at a profound informational disadvantage. This suggests that the democratization of these tools (through accessible DMaaS) is not just a market opportunity but a regulatory necessity to prevent monopolistic information dominance.

4.8 Latency Reduction in Decision Cycles

One of the most profound impacts of integrating Adaptive Data Engineering is the drastic reduction in the "Data-to-Decision" latency. In traditional environments, the cycle from data generation (e.g., a sales transaction or a market event) to its availability in a BI dashboard can take 24 to 48 hours due to batch processing windows.

LLM-ADE enables "Micro-Batching" or near-real-time streaming. Because the schema transformation is handled by the AI model dynamically, the rigid "staging area" requirements of traditional data warehouses are minimized. This allows executives to react to market shifts within minutes. For example, if an LMM detects a surge in negative sentiment regarding a product on social media (Visual/Textual), the BI system can immediately trigger an alert to the supply chain team to adjust inventory orders, potentially saving millions in overstock costs. This agility is a direct realization of the "long-term competitiveness through IT assets" predicted by Ross et al. [4] nearly three decades ago.

5. DISCUSSION

5.1 Strategic Implications for Enterprise

The integration of LLM-ADE and LMMs suggests that the "moat" of a business is no longer its data volume, but its data fluidity. Companies that cling to rigid, siloed data structures will find themselves outmaneuvered by competitors who can reconfigure their informational assets on the fly. This validates the TOE framework's emphasis on "Technological Readiness" [6]; readiness is now defined by adaptability rather than capacity.

5.2 The Risk of Hallucination in Business Intelligence

A critical limitation of this new architecture is the propensity for Generative AI to hallucinate. In a creative context (like the "ChatGPT as Inventor" study [11]), a hallucination might be a spark of novelty. In a financial reporting or competitive bidding context [3], a hallucination could be catastrophic.

If an LLM "invents" a competitor's weakness that does not exist, and the firm bases a strategic bid on that false premise, the financial losses could be substantial. Therefore, the implementation of "Retrieval-Augmented Generation" (RAG) within BI systems is mandatory. RAG forces the LLM to ground its answers in retrieved, verifiable documents, reducing the hallucination rate. However, even with RAG, the "Trust Gap" remains a significant barrier to widespread adoption.

5.3 Privacy and Proprietary Information

The use of third-party LLMs (like GPT-4 via API) to process proprietary business data raises immense security concerns. As Richard et al. [3] discuss, proprietary information is the lifeblood of competitive bidding. Sending this data to a cloud provider for processing constitutes a risk. This drives a growing trend toward "Small Language Models" (SLMs) or open-source models (like Llama 3 or Mistral) hosted within the enterprise's private cloud (VPC). This allows the benefits of intelligent automation without the risk of data leakage.

5.4 Future Directions: The Autonomous Enterprise

Looking forward, the convergence of these technologies points toward the "Autonomous Enterprise." In this theoretical end-state, the BI system does not just report on the business or predict the future, but actively intervenes. An advanced LLM-ADE system could theoretically detect a supply chain disruption, identify an alternative supplier, negotiate pricing (within set parameters), and execute the purchase order—all without human intervention. While this remains speculative, the building blocks—automated ETL [13], generative reasoning [11], and multimodal sensing [9]—are already in place.

6. CONCLUSION

The transition from traditional Business Intelligence to Intelligent Automation is not merely a technical upgrade; it is an organizational metamorphosis. The synthesis of Adaptive Data Engineering (LLM-ADE) with Large Multimodal Models (LMMs) resolves the historic bottleneck of data preparation, allowing organizations to process the chaotic reality of the market into structured strategic insights.

Our analysis confirms that while the barriers to adoption for SMEs remain high [5, 7], the cost of inaction is higher. As tech giants [10] leverage these tools to dominate information landscapes, the market will increasingly bifurcate into those who use AI to navigate uncertainty and those who are overwhelmed by it.

Future research must focus on the governance frameworks required to manage these autonomous systems. As we hand over more of the "intelligence" in Business Intelligence to algorithms, we must ensure that the "Business" objectives—profitability, ethics, and sustainability—remain firmly in human hands.

REFERENCES

1. Jermol, M., Lavrac, N., & Urbancic, T. (2003). Managing business intelligence in a virtual enterprise: A case study and knowledge management lessons learned. Journal of Intelligent & Fuzzy Systems, 14(3), 121-136.
2. Oyku, I., Mary, C. J., & Anna, S. (2012). Business intelligence success: The roles of BI capabilities and decision environments. Information & Management, 50, 13-23.
3. Richard, E. W., Paul, R. M., & Robert, J. W. (n.d.). Competitive Bidding and Proprietary Information. Journal of Mathematical Economics, 11, 161-169.
4. Ross, J. W., Beath, C. M., & Goodhue, D. L. (1996). Develop long-term competitiveness through IT assets. Sloan Management Review, 38(1), 31–44.
5. Zide, O., & Jokonya, O. (2021). Factors affecting the adoption of Data Management as a Service (DMaaS) in Small and Medium Enterprises (SMEs). Procedia Computer Science, 196, 340–347.
6. Basloom, R. S., Sani Mohamad, M. H., & Auzair, S. M. (2022). Applicability of public sector reform initiatives of the Yemeni government from the integrated TOE-DOI framework. International Journal of Innovation Studies, 6(4), 286–302.
7. Wessels, T., & Jokonya, O. (2021). Factors affecting the adoption of big data as a service in SMEs.

Procedia Computer Science, 196, 332–339.

8. Härting, R. C., & Sprengel, A. (n.d.). Cost-benefit considerations for data analytics.
9. Fan, L., Lee, C.-H., Su, H., Feng, S., Jiang, Z., & Sun, Z. (2024). A New Era in Human Factors Engineering: A Survey of the Applications and Prospects of Large Multimodal Models. arXiv preprint arXiv:2405.13426.
10. Patel, D. B. (2025). Leveraging BI for Competitive Advantage: Case Studies from Tech Giants. Frontiers in Emerging Engineering & Technologies, 2(04), 15–21.
11. Ege, D. N., Øvrebø, H. H., Stubberud, V., Berg, M. F., Elverum, C., Steinert, M., & Vestad, H. (2024). ChatGPT as an inventor: Eliciting the strengths and weaknesses of current large language models against humans in engineering design. arXiv preprint arXiv:2404.18479.
12. Choi, S., & Gazeley, W. (2024). When Life Gives You LLMs, Make LLM-ADE: Large Language Models with Adaptive Data Engineering. arXiv preprint arXiv:2404.13028.
13. Mantri, A. (2024). Intelligent Automation of ETL Processes for LLM Deployment: A Comparative Study of Dataverse and TPOT. European Journal of Advances in Engineering Technology, 11, 154–158.

.