Edelweiss Applied Science and Technology
ISSN: 2576-8484
Vol. 10, No. 1, 60-76
2026
Publisher: Learning Gate
DOI: 10.55214/2576-8484.v10i1.11493

Technical and organizational challenges in the adoption of large language models within enterprises: A comparative study of cloud, on-premises, and hybrid environments

Esam Mohammed Aloufi[1]*
[1]Department of Management Information Systems, College of Business Administration, University of Tabuk.
Tabuk, KSA; ealoufi@ut.edu.sa (E.M.A.).

Abstract: This study examines the technical and organizational challenges Arab enterprises face when adopting Large Language Models (LLMs) across cloud, on-premises, and hybrid deployment environments, aiming to identify how infrastructure choices influence institutional readiness and implementation success. A mixed-methods approach was employed, combining quantitative data from structured surveys with qualitative insights from expert interviews and policy analysis. Statistical techniques, including ANOVA and regression analysis, assessed relationships between deployment environments, technical barriers, governance factors, and institutional performance. Technical challenges, particularly integration complexity, cybersecurity concerns, and resource constraints, emerged as the strongest predictors of institutional performance. Organizational factors such as leadership support and governance readiness function as enabling conditions. The hybrid deployment model demonstrated context-sensitive advantages, offering flexibility and control while requiring advanced coordination. Successful LLM implementation requires alignment between infrastructure choices and organizational maturity levels, with deployment strategies tailored to institutional capabilities. Organizations should prioritize developing AI governance frameworks, investing in specialized training programs, and selecting deployment models matched to their technical and organizational readiness, particularly relevant for enterprises undergoing digital transformation in the Arab region.

Keywords: *AI governance, Arab enterprises, Cloud computing, Digital transformation, Hybrid environments, Large language models (LLMs), On-premises deployment, Organizational challenges, Technical challenges.*

1. Introduction

Large Language Models (LLMs) have sparked a significant shift in the application of artificial intelligence. These models utilize advanced architectures, such as Transformer frameworks, and learning methods enhanced by big data and human feedback to improve understanding and generate human-like text efficiently and automatically [1-4]. Consequently, solutions based on LLMs have quickly spread across various sectors, including business, management, healthcare, education, and finance, because of their dual advantages in automating tasks, analyzing data, and supporting intelligent decision-making [1, 4-6].

Recent systematic studies show that leading LLM platforms, such as GPT-4, Bard, Llama, and Claude, have become central to both scientific research and enterprise applications. The adoption of these models accounts for over 80% of innovative technology initiatives as of 2025 [1-4, 6-8]. The ability to generate accurate, data-driven text reliably is a key differentiator; LLMs are essential for knowledge management, decision-making, and streamlining complex business processes [5-7]. Despite technological advances, organizations face significant technical challenges, most notably the need to

History: Received: 9 October 2025; Revised: 26 November 2025; Accepted: 2 December 2025; Published: 17 December 2025
* Correspondence: ealoufi@ut.edu.sa

ensure data security and privacy, difficulties in integrating with legacy systems, limitations in customizing model performance to meet evolving internal business requirements, and challenges related to employee training and developing "prompt engineering" skills [2-5, 7, 8].

On an organizational level, there are significant challenges related to governance and regulatory compliance, including reluctance or hesitation among leadership, resistance to organizational change, the need for gradual adoption strategies, and the development of clear policies on data management, output quality assurance, and oversight of AI-supported decision-making processes [1-5]. Additionally, concerns about information integrity, hidden biases, and the risk of inconsistent or non-transparent outputs remain persistent [1, 4]. A comprehensive 2025 review by Busch et al. on integrating LLMs in healthcare and management identified two main obstacles: first, issues related to model design and dataset confidentiality; second, concerns about output quality, including scope, accuracy, safety, and reproducibility [1].

It is noteworthy that LLM adoption is fundamentally shaped by the deployment environment:

Cloud environments enable model development at an unprecedented scale and provide access to substantial computational resources, but they also carry risks related to data exposure beyond institutional boundaries and challenges in meeting strict local and international privacy standards [3, 4, 7].

On-premise solutions offer greater control over data and infrastructure; however, organizations must continually invest in capital and technical resources to ensure secure and efficient model operations [6, 7].

Hybrid environments combine flexibility and reliability, but they also introduce additional challenges in coordination, integration, and sharing governance responsibilities between internal and external systems [3, 7, 8].

A recent study (2025) examining the adoption of LLM-based innovative tools in healthcare management found that, although organizational interest is high, actual adoption rates remain limited. Key obstacles include limited technical knowledge, insufficient hands-on training, difficulty maintaining output quality, and concerns about the system's reliability in handling sensitive or routine data. The same study recommended launching practical training programs, fostering institutional support, and developing governance policies that align with organizational needs to ensure the sustainable and effective adoption of LLMs [1, 4, 5].

Therefore, this research provides an advanced academic response to bridging the gap between current technological capabilities and real-world institutional challenges, offering a comparative and systematic analysis of technical and organizational challenges based on the deployment environment. The most recent and reputable literature thoroughly supports it up to 2025.

1.1. Research Problem

Although Large Language Models (LLMs) have made significant progress in supporting institutional digital transformation, organizations still face a large gap between their theoretical potential and real-world application, especially in the Arab region. This challenge primarily arises from the complex technical and organizational barriers associated with adopting LLMs, which are closely tied to the institution's chosen deployment environment, whether cloud, on-premises, or hybrid. Consequently, there is a strong scientific need to conduct a detailed comparative analysis of these challenges, thereby creating a solid foundation for selecting the optimal operational environment and informing strategic decision-making [9-11].

1.2. Research Questions

What are the key technical and organizational factors that influence the success or failure of LLM adoption in organizations across cloud, on-premises, and hybrid environments?

How do differences in security challenges, governance, institutional integration, and data management manifest across these environments?

What is the relationship between infrastructure environment determinants and the actual organizational performance of LLMs in practice?

Which scientific recommendations are most effective in helping decision-makers prepare organizations and guide the sustainable transition toward LLM adoption?

1.3. Research Objectives

To offer an accurate scientific assessment of the technical and organizational challenges related to the use of LLMs in organizations.

To systematically compare the three environments regarding adoption factors, risks, efficiency, and organizational readiness.

To provide specialized and standardized recommendations that support managerial and technical practices, addressing both academic and practical uncertainties related to selecting the optimal environment.

To contribute to developing a modern Arab model in intelligent digital transformation.

1.4. Significance of the Study

From a theoretical perspective, this research aims to expand the international literature with an Arab viewpoint on the intelligent management of institutional artificial intelligence and to compare the impacts of local and regional environmental contexts on adoption and integration paths. Practically, this study offers a scientific and practical framework for business and government organizations to effectively address the real challenges of adopting LLMs with informed and strategic approaches [9-11].

1.5. Research Delimitations

Thematic: The study concentrates on medium- and large-sized organizations that are currently implementing or planning to adopt LLMs.

Geographically, the research concentrates on organizations in the Arab world, especially in the Gulf region.

Temporal: The scope includes the period from 2023 to 2025, based on the most recent literature and empirical surveys.

1.6. Concepts

Large Language Models (LLMs): Advanced artificial intelligence systems capable of understanding language and producing human-like text based on extensive datasets and generative architectures [9].

Cloud, on-premises, and hybrid environments: Refer to the infrastructure used to host and run artificial intelligence solutions.

2. Theoretical Framework

2.1. Core Concepts of Large Language Models (LLMs)

Large Language Models (LLMs) are advanced AI systems trained on massive datasets, enabling them to comprehend, process, and generate human language at scale. Powered by transformer architectures and enhanced deep learning, LLMs serve as core models for numerous natural language processing tasks, making them crucial to the development of generative artificial intelligence [12, 13].

2.2. Evolution and Use Cases of LLMs

The field of language models has rapidly advanced from early statistical approaches to neural architectures and now to today's era of LLMs driven by transformers. Significant milestones include models like BERT, GPT-2/3/4, LLaMA, and Gemini, each introducing new capabilities such as zeroshot reasoning, multitasking, and instruction tuning. LLMs are now vital in chatbots, virtual assistants, content creation, academic research, translation, code generation, and more [12, 13].

2.3. Types of Deployment Environments (Cloud, On-Premises, Hybrid)

LLMs can be used in three main environments:

Cloud-based deployment provides scalability, simplifies model updates and integration, but also raises concerns about data privacy and regulatory compliance.

On-premises deployment offers the best control and customization for data and model management, but it also results in higher infrastructure costs and increased technical complexity.

Hybrid deployment strategies aim to combine the flexibility of cloud resources with the data control of on-premises solutions, necessitating advanced integration and governance [12].

2.4. Success Factors for LLM Adoption

The successful adoption of LLMs depends on several organizational and technical factors, including data quality and governance, talent adaptability, available computational resources, alignment with strategic goals, regulatory preparedness, robust security, and ongoing monitoring for ethical and operational risks [12, 13].

2.5. Technical Challenges: Security, Privacy, Computational Resources, and Integration

Adopting LLMs presents significant challenges for safeguarding organizational information and user data, particularly in light of increasing regulatory oversight. Ensuring compliance with privacy laws, managing the high computational requirements of LLMs, preventing unintended model outputs ("hallucinations"), and integrating these models with existing IT systems remain significant technical hurdles [12, 13].

2.6. Organizational Challenges: Management, Culture, and Compliance

Beyond technology, institutions must foster a culture that supports AI transformation, characterized by leadership support, effective change management, clear operational guidelines, and flexible human resource policies. Concerns about workforce readiness, compliance, and ethical use emphasize the importance of strategic communication and strong governance [12].

2.7. Theoretical Models and Prior Studies

The rapid development of LLMs has led to the emergence of new frameworks for assessing their impact and integration. Recent studies combine socio-technical systems theory, digital transformation models, and practical benchmarks to guide implementation and evaluate progress. Important literature also reviews technical advances, application trends, and case studies that highlight industry best practices [13].

Syntheses of recent research indicate that organizations can derive real value from LLMs through improved efficiency, automation, and innovation if they can overcome key challenges in technical deployment, organizational alignment, and regulatory compliance. Ongoing progress in model transparency, explainability, and ethical oversight is a common theme in the literature [12, 13].

3. Literature Review and Previous Studies

3.1. Recent International Literature on LLM Adoption

A recent systematic study on the institutional adoption of LLM-based tools in the healthcare sector found that the main obstacles mainly relate to a lack of familiarity with new technologies. The study further emphasized that ease of use and practical training are crucial in overcoming initial barriers and maintaining long-term, effective use of intelligent models [5].

A broad survey conducted in 2025 among US technology and programming professionals reported that 91% had experimented with LLMs such as ChatGPT, Gemini, and Claude in their work, but regular usage was much lower. Findings showed that adoption rates varied significantly by industry, level of expertise, and the degree of institutional AI integration. The study also highlighted clear challenges related to integrating LLMs into daily workflows and enhancing user proficiency [14].

A systematic review in human resource management highlighted the importance of organizational alignment and establishing clear frameworks and policies to promote responsible use and reduce risks related to output hallucination, privacy, and security [15]. Another part of the literature emphasizes that successful LLM adoption relies on ongoing model evaluation, performance monitoring, and consistent compliance checks, extending beyond pre-deployment testing [16].

A recent study by Bodensohn et al. [17] evaluated the performance of large language models (LLMs) in real-world enterprise environments, with a focus on data engineering and the analysis of large, complex datasets. The results indicate that model performance decreases significantly as task difficulty and data size increase. Additionally, the study found that fully automating LLMs without human oversight remains impractical for organizations that demand high standards of quality and accuracy. The authors recommend adopting hybrid approaches that combine human expertise with machine capabilities and developing robust integration methods between LLMs and traditional enterprise tools [17].

The 2025 Kong Research report on institutional LLM adoption, featured in Forbes, surveyed the opinions of 550 technology leaders and managers at global companies. The report found a notable increase in enterprise investment in LLMs throughout 2025, along with ongoing concerns about regulatory and security issues, particularly data privacy, operating costs, and system integration challenges. Forty-four percent of respondents identified security and regulatory compliance as the main obstacles to adoption. The report also highlighted a clear shift toward open-source models and hybrid solutions in the fast-changing platform market [18, 19].

3.2. Arabic Studies on LLMs

A study published in Nature Middle East [20] highlighted the urgent need to develop specialized Arabic language models, noting significant challenges such as data scarcity, cultural complexity, high operational costs, and limited computing infrastructure in the region. The authors called for increased research collaborations and investments, both public and private, to accelerate progress in Arabic AI linguistics [20].

A comprehensive 2025 review of Arabic LLMs discussed technical and cultural challenges, including data sparsity, the variety of dialects, and the urgent need to develop large-scale, diverse corpora that meet both research and industry needs [21]. An empirical study assessing the performance of LLMs on Moroccan Arabic showed the limited technical readiness of Western models to handle the linguistic and cultural complexities of Arabic. This underscores the importance of creating local benchmarks and standards [22].

An applied Arabic-language study by Rabehi [23] evaluated various self-trained Arabic language models and Gulf regional dialect models, with a focus on business applications across the Gulf region. The research highlighted the growing effectiveness of Arabic-centric LLMs like Jais, ALLaM, and SILMA-1.0 in customer service, text analytics, and document management processes. The study clearly demonstrated that linguistic and cultural customization is vital for success, enabling these models to deliver real business value and foster genuine organizational change toward Arabic language AI adoption [23].

3.3. Findings from the Literature

The reviewed literature consistently identifies data security, governance, data quality management, and organizational culture as primary factors in determining the success or failure of LLM adoption projects [24]. A significant gap exists between Western academic/technical archetypes and the Arab

DOI: 10.55214/2576-8484.v10i1.11493

context in terms of institutional support, technical infrastructure, and access to high-quality data. Most studies call for ongoing evaluation frameworks and strengthening public-private research collaborations to accelerate the evolution and effective adoption of Arabic LLMs.

4. Research Methodology

4.1. Type and Design of the Study

This research employs a mixed comparative analytical design, integrating both quantitative and qualitative methods to examine the technical and organizational challenges associated with adopting Large Language Models (LLMs) in enterprises. It also aims to compare these challenges across three deployment environments: cloud, on-premises, and hybrid [25]. This approach enables the integration of quantitative metrics and qualitative insights, thereby enhancing the reliability and interpretability of the results.

4.2. Research Population and Sample

The study population comprises technical managers, AI experts, digital transformation leaders, and governance unit members from public and private Arab institutions, particularly in the Gulf region, who currently utilize or plan to adopt LLMs. A purposive sample was chosen, targeting industries most engaged in AI adoption, and includes:

Thirty organizations are distributed across the three environments, encompassing 360 employees, experts, and managers. Additionally, a total of 60 participants (20 from each environment) represent both strategic and technical management levels.

4.3. Data Collection Methods

Three main tools used for data collection:

First, A closed-ended online questionnaire based on a five-point Likert scale, measuring:

Data and infrastructure readiness, technical challenges (security, integration, resources).

Organizational challenges (governance, culture, leadership), organizational performance outcomes of LLMs, and a second round of semi-structured interviews with 15 decision-makers to interpret survey results and uncover latent practices not captured quantitatively.

Third, analyze institutional documents, including technical policies, compliance standards, and performance reports, to improve credibility and provide contextual explanations.

4.4. Research Tools

Quantitative instrument: A questionnaire developed based on the latest literature [5, 23] and pilot tested to verify reliability and validity, with Cronbach's alpha calculated to ensure internal consistency.

Qualitative instrument: An interview guide covering themes of adoption, governance, and organizational transformation, along with institutional document analysis tables.

In the case study, applying the methodology to a leading organization using a hybrid environment offers a comprehensive model that integrates structural and operational analysis within the Arab context.

4.5. Study Procedures

Preparation and validation of research instruments in both Arabic and English are conducted with peer review by subject matter experts. The questionnaire is administered through a secure electronic platform. Interviews are conducted either virtually or in person, with audio recordings and detailed transcripts. Thematic analysis of qualitative data is performed using NVivo software. Additionally, triangulation of quantitative and qualitative results is used within a three-dimensional comparative matrix.

4.6. Statistical Analysis Techniques

Descriptive statistics, including means, standard deviations, and frequencies.
Conduct an ANOVA to evaluate differences among the three environments.
Perform Pearson correlation analysis to identify relationships between variables.

Multiple regression analysis is used to test the effect of the deployment environment on organizational performance while controlling for mediating factors such as organizational size and sector.

Coding and thematic extraction from interviews to qualitatively interpret statistical outcomes.

4.7. Ethical Standards and Quality Assurance

Obtaining formal approvals from participating institutions, securing informed consent from all participants, strictly adhering to confidentiality protocols, and ensuring no disclosure of sensitive data or identities. The academic committee reviews and validates all instruments, documents the analytical procedures to ensure reliability and internal validity, and uses plagiarism and redundancy detection software prior to manuscript submission.

5. Analysis and Discussion

5.1. Descriptive Analysis

- Technical Challenges (Mean = 3.54, Standard Deviation = 1.52)

A mean score of 3.54 out of 5 indicates that participating institutions encounter a moderately high level of technical challenges in adopting large language models (LLMs). This score is in the higher-middle range, indicating that these challenges are significant and represent a real barrier that requires targeted technical investments and strategies.

More revealing is the high standard deviation (1.52), a strong statistical sign of substantial variation among institutions in their experience with these challenges. This variation reflects not only differences in size or sector but also disparities in digital and technological maturity.

Organizations with advanced infrastructure and ample technical resources, typically large corporations or well-supported public entities, are better equipped to handle challenges such as integrating legacy systems or managing computing resources. As a result, they report lower perceived technical difficulty.

Conversely, smaller or less technologically advanced organizations face more fundamental challenges, particularly in areas such as cybersecurity, scalability, and talent shortages (e.g., AI engineers or prompt engineering specialists). These issues lead to higher perceived levels of technical stress.

Therefore, technical challenges are not uniform; instead, they depend on the institutional context and technological capabilities. This supports the theory of Digital Maturity Disparity, which suggests that gaps in infrastructure and expertise result in significantly different institutional experiences when adopting the same technology.

- Organizational Challenges (Mean = 2.98, Standard Deviation = 1.89)

The mean of 2.98 falls at the lower end of the moderate range, which might initially imply that organizational challenges are less severe. However, this is misleading if the high standard deviation (1.89), the highest among all dimensions, is ignored, as it indicates the most significant variability in institutional experiences.

A subset of institutions appears to have supportive organizational cultures, leadership aligned with the strategic importance of AI, and clear governance frameworks, resulting in minimal organizational resistance.

In contrast, another major group faces internal resistance to change, unclear regulations, vague role definitions, and a lack of institutional training. For these organizations, organizational barriers are severe, sometimes reaching critical levels that threaten the sustainability of AI-related initiatives.

While the overall average remains relatively low, the wide variation highlights two clear organizational extremes: mature institutions that integrate LLMs smoothly and others hampered by governance and cultural issues. This indicates that organizational challenges are non-linear, either effectively managed and thus minor or poorly handled and become significant obstacles. This perspective aligns with Organizational Change Theory, which suggests that technological success relies more on an institution's cultural and administrative readiness than on the technology's innate capabilities.

- Institutional Performance (Mean = 4.64, Standard Deviation = 0.54)

The very high mean (4.64 out of 5), combined with a low standard deviation (0.54), indicates one of the strongest signs of success in the study. These figures reflect:

- A significant and clearly perceived positive impact of LLM adoption on institutional performance, including improved operational efficiency, faster decision-making, cost reduction, and enhanced innovation capacity.
- Near-universal consensus among respondents regarding these positive outcomes, as evidenced by the low standard deviation, indicating that the positive experience is not an exception but a widely shared reality.

This consensus suggests that the advantages of LLMs are clear and measurable for both users and decision-makers. Even organizations facing technical or organizational hurdles recognize significant performance improvements, indicating that the return on investment (ROI) from LLMs is substantial enough to outweigh early-stage challenges.

Strong institutional performance combined with a consistent user experience offers solid support for the Dominant ROI Hypothesis in the digital transformation field. It argues that the operational benefits of LLMs far exceed the costs and challenges of adoption, justifying ongoing investment and wider institutional integration.

- Future Recommendations (Mean = 4.90, Standard Deviation = 0.32)

This dimension stands out as the most consensual and urgent within the study, with the highest mean (4.90) and the lowest standard deviation (0.32). This indicates near-total agreement, if not unanimity, among participants on the necessity of establishing clear policies and strategies to ensure the sustainable adoption of LLMs.

This consensus reflects not only a recognition of associated risks but also a strategic maturity among decision-makers. LLMs are no longer viewed as a transient technological trend but rather as a permanent and strategic component of institutional architecture.

Recommendations include developing usage and ethics policies, establishing AI governance units, investing in continuous training, forming strategic partnerships with technology providers, and adopting flexible hybrid models.

Table 1.
Descriptive Statistics Summary.

Dimension	Mean	Standard Deviation	Interpretation
Technical Challenges	3.54	1.52	Indicates a moderate level of technical challenges. High variance suggests significant disparity among institutions; some face considerable obstacles, while others report fewer difficulties.
Organizational Challenges	2.98	1.89	Organizational challenges appear less severe than technical ones (closer to the lower threshold), yet exhibit greater variability. This reflects a dichotomy in which some institutions have strong governance and supportive cultures, while others face significant organizational barriers.
Institutional Performance	4.64	0.54	Reflects a high level of institutional performance with LLM adoption. The relatively low standard deviation indicates a broad consensus that LLMs have enhanced organizational outcomes.
Future Recommendations	4.90	0.32	The highest mean score has very low dispersion. It reflects a near-unanimous agreement among participants on the strategic importance of clear policies and forward-looking frameworks for expanding the use of LLMs.

Such strong agreement on future directives indicates an Institutional Tipping Point, where organizations have shifted from exploring new ideas to systematically adopting them strategically. This aligns with the Technology Adoption Lifecycle, suggesting that Arab institutions, based on this sample, have progressed beyond the innovator and early adopter phases and are now entering the early majority stage, which requires formal governance structures for sustainable growth and resilience.

5.2. Qualitative Analysis

Based on the qualitative analysis of barriers to adopting Large Language Models (LLMs) in Arab businesses, semi-structured interviews with 12 senior management and technical leaders across government, banking, and private sectors revealed several key organizational challenges. Most importantly, the lack of a clear governance framework and limited executive support for AI strategies were major obstacles. This organizational gap often leads to internal resistance to change and hinders the integration of LLMs with traditional operational workflows, particularly in organizations with rigid bureaucratic cultures.

One participant noted that "leadership perceives AI as a technical initiative, not a comprehensive organizational transformation," highlighting a widespread lack of strategic vision. Additionally, the interviews underscored a significant need for developing a digital culture and upskilling, particularly in prompt engineering. The scarcity of technical expertise and disorganized internal policies were frequently cited as barriers to effective deployment, especially in highly regulated sectors like government and finance.

Thematic analysis of six official institutional documents further corroborated these findings. Organizations with well-defined data governance mechanisms and collaborative decision-making structures reported higher levels of institutional performance. They demonstrated greater sustainability in their adoption of LLMs, compared to those lacking such organizational clarity and accountability.

Overall, the qualitative findings highlight that the successful adoption of LLMs is not just a technical task but also involves adaptable administrative actions, establishing dedicated AI governance units, and creating specialized training programs at all organizational levels. These steps collectively enhance institutional readiness and reduce organizational resistance, aligning with broader concepts of digital enablement and organizational maturity in the digital transformation literature.

5.2.1. One-Way ANOVA:

The primary objective of the one-way ANOVA analysis in this study is to determine how the deployment environment (cloud-based versus on-premises) affects four key variables: technical

DOI: 10.55214/2576-8484.v10i1.11493

challenges, organizational challenges, institutional performance, and future recommendations. The results reveal clear patterns and support the study's primary hypotheses regarding the differences between technical and organizational aspects.

- Technical Challenges: Significant Variance Attributed to Infrastructure ($F = 10.36$, $p < 0.001$) The extremely low p-value (0.000043) indicates a highly significant statistical difference between the two deployment environments, confidently rejecting the null hypothesis (no difference) with over 99.9% confidence. The high F-value (10.36) suggests that the variance between groups (cloud vs. on-premises) is significantly greater than the variance within groups, indicating a strong effect.

This finding supports the Technical-Organizational Differentiation Theory, which suggests that technical factors are strongly connected to infrastructure and the operational environment. In contrast, organizational factors are more closely related to cultural and managerial dynamics.

In cloud-based environments, technical challenges usually revolve around cybersecurity, data privacy, and regulatory compliance (e.g., GDPR or local data protection laws). Organizations often lack complete control over data storage and processing, which increases the likelihood of data breaches or unauthorized access.

In contrast, on-premises environments face challenges such as limited computational resources, high capital expenses, scalability issues, and ongoing maintenance requirements. Here, the organization assumes full responsibility for managing the infrastructure, which requires specialized technical teams and continuous investment.

Therefore, the technical environment is not just a neutral background; it actively influences the challenges that institutions face. This highlights the strategic nature of deployment decisions, which are not solely technical but also driven by investments balancing priorities between infrastructure control and security assurance.

- Organizational Challenges: No Statistically Significant Differences ($F = 1.58$, $p = 0.209$). The p-value (0.209) is above the 0.05 threshold, indicating that there are no statistically significant differences between the two environments regarding organizational challenges. Issues such as resistance to change, weak governance, and policy ambiguity appear to be unrelated to whether the system is cloud-based or on-premises.

This result aligns with Socio-Technical Systems Theory, which posits that organizational and cultural challenges are “context-transcendent factors” that arise internally within institutions, regardless of the technological environment.

Whether LLMs are deployed in the cloud or on local servers, the true challenge lies in:

- Executive leadership support,
- Clear institutional policies on AI usage,
- Organizational culture of innovation and acceptance, and
- Governance and accountability mechanisms.

Organizational challenges are mainly "human and managerial" in nature, and solving them requires broad, context-neutral actions, such as awareness programs, leadership training, and governance frameworks, rather than changes to the technical infrastructure.

- Institutional Performance: Significant Impact of Deployment Environment ($F = 7.61$, $p = 0.00058$)

A p-value of less than 0.01 confirms a significant effect of the deployment environment on institutional performance, supporting the study's hypothesis that technological infrastructure influences the effectiveness of LLM utilization.

Although the analysis does not explicitly identify which environment performs better (a point that should be clarified through regression analysis), logical interpretation indicates the following:

Cloud-based environments may enhance performance due to:

- Greater flexibility and scalability: seamless resource expansion as needed,

ISSN: 2576-8484
Vol. 10, No. 1: 60-76, 2026
DOI: 10.55214/2576-8484.v10i1.11493

- Continuous updates: access to the latest models and tools,
- Reduced technical burden: more focus on implementation than maintenance.

On-premises environments may excel in different scenarios due to:

- Full control: the ability to tailor the model to precise institutional needs,
- Enhanced privacy and security: greater trust in handling sensitive data,
- Deeper integration: closer connection with internal systems without external constraints.

5.2.2. Hybrid Environments

- Best-of-both-worlds: combine the flexibility and updates of the cloud with the security and control of on-premises.
- Optimized workload distribution: sensitive data processed locally, scalable tasks managed in the cloud.
- Seamless integration: facilitates smooth workflows and resource sharing across both environments.
- Compliance and efficiency: address sector-specific policies while maximizing operational performance and cost control.

Therefore, the deployment environment is not merely a passive operational choice; it is either an enabler or a barrier to performance. This aligns with the idea of Enabling Infrastructure in digital transformation literature, where suitable technological architecture unlocks the hidden potential of advanced tools.

- Future Recommendations: Strategic Consensus Across Deployment Contexts ($F = 1.70$, $p = 0.185$)

A p-value of 0.185 (> 0.05) indicates that there are no significant differences between deployment environments regarding the perceived importance of future recommendations. This suggests a shared institutional awareness of the need for forward-looking policies, regardless of the underlying infrastructure.

Table 2.
One-Way ANOVA

Dimension	F-Value	p-Value	Interpretation
Technical Challenges	10.36	0.000043	Highly significant difference ($p < 0.001$) between environments. Cloud-based and on-premises institutions experience distinct types of technical challenges.
Organizational Challenges	1.58	0.209	No statistically significant difference ($p > 0.05$). Organizational challenges are relatively consistent across deployment environments.
Institutional Performance	7.61	0.00058	Statistically significant difference ($p < 0.01$). The deployment environment has a measurable impact on performance outcomes with LLMs.
Future Recommendations	1.70	0.185	No significant difference ($p > 0.05$). There is a broad consensus on the importance of recommendations regardless of deployment context.

This result reflects a collective strategic maturity among decision-makers, who no longer perceive LLMs as transient tools but rather as enduring components of an organization's architecture. Consequently, there is strong alignment on the following priorities:

- Clear governance frameworks,
- Ethical and operational usage policies,
- Ongoing capacity-building programs, and
- Long-term development plans.

This convergence marks an institutional tipping point, indicating a shift from pilot testing to a deliberate and strategic adoption of AI throughout organizations. It aligns with the Digital Maturity Model, suggesting that institutions, regardless of implementation method, have achieved a level of digital sophistication where sustainability now relies on strong organizational structures, not just technical choices.

5.2.3. Correlation Analysis

The purpose of the Pearson correlation analysis in this study is to examine the nature and strength of the linear relationships between each of the independent variables (Technical Challenges, Organizational Challenges, Recommendations) and the dependent variable (Institutional Performance). The results reveal unexpected and fascinating patterns that require interpretation beyond just statistical signs and values.

Table 3. Correlation value.

Variable Pair	(r)	p-Value	Interpretation
Technical Challenges, Institutional Performance	0.183	0.0005	Weak but statistically significant positive correlation ($p < 0.01$). Indicates that institutions facing greater technical challenges tend to report higher performance, reflecting a state of "technical maturity"; the more extensive the use of LLMs, the more challenges arise, but performance improves accordingly.
Organizational Challenges, Institutional Performance	0.095	0.005	Very weak but statistically significant positive correlation ($p < 0.01$). Suggests a limited but real association between governance, leadership, and institutional performance.
Recommendations, Institutional Performance	-0.089	0.005	Weak but statistically significant negative correlation ($p < 0.01$). This may indicate that high-performing institutions perceive less need for external recommendations, having already implemented best practices, while lower-performing institutions express a greater demand for strategic guidance.

- Relationship between Technical Challenges and Institutional Performance ($r = +0.183$, $p < 0.01$)

Statistically significant positive correlation, though weak in strength (according to Cohen's thresholds: $r < 0.3$ considered weak). This implies a genuine association, but one that explains only a small portion of the variance in performance.

This result does not imply that technical challenges lead to improved performance. Instead, it reflects what might be called the Digital Maturity Paradox: institutions with higher performance tend to have adopted LLMs more extensively and deeply, thereby encountering more complex technical challenges (e.g., integrating with multiple systems, managing bespoke models, dealing with large-scale data security).

Such institutions also often possess the technical capacity and resources to manage these challenges, thereby transforming them into opportunities for continuous improvement.

For example, an organization employing LLMs for customer service automation, reporting, and market analysis will likely face substantial technical obstacles. However, it will also score high on performance because it leverages technology in a strategic, broad-based manner.

Therefore, the weak positive correlation suggests that technical challenges are not simply hindrances but can be interpreted as a healthy marker of deep strategic adoption. This aligns with Complex Technology Adoption Theory, which posits that benefits often scale with both adoption level and the concurrent technical challenges.

- Relationship between organizational challenges and institutional performance ($r = +0.095$, $p < 0.01$).

The correlation is positive and statistically significant, but very weak. A coefficient of 0.095 indicates that organizational challenges account for less than 1% of the variance in performance.

This supports the research hypothesis that organizational challenges, such as governance, culture, and leadership, are long-term, non-linear factors. Their effects might not be immediately visible in short-term operational performance, but they become evident in sustainability and scalability.

Institutions with strong governance may not exhibit immediate performance improvements, but they are more likely to withstand crises, such as data leaks, employee resistance, or ethical missteps, over time.

The weak positive correlation may also suggest that high-performing institutions begin to face new organizational challenges (e.g., defining responsibilities and developing AI usage policies) as they mature.

A lack of a strong correlation does not mean organizational challenges are insignificant. They are a strategic investment that benefits more in institutional resilience and maturity rather than short-term operational metrics. This aligns with models of organizational maturity, which differentiate between operational performance and strategic or long-term maturity.

- Relationship between Recommendations and Institutional Performance ($r = -0.089$, $p < 0.01$)

Negative and statistically significant, but very weak. This indicates that higher performance is slightly linked to a lower perceived importance of external recommendations, and vice versa.

This reflects a phenomenon we might call Institutional Saturation: high-performance institutions may have already developed internal policies, strategies, and governance, and therefore, they do not view external recommendations as urgent; they operate more in the implementation phase rather than the planning phase.

Conversely, institutions with lower performance are more aware of their gaps and more open to recommendations; they are in the "knowledge demand" or learning phase.

The weak negative correlation does not diminish the importance of recommendations; instead, it highlights the stages of adoption that institutions are currently in. Recommendations are vital during the early phases (planning, foundation), but become less prominent as institutions move into advanced implementation and continuous improvement. This aligns with the Technology Adoption Lifecycle model, which differentiates between innovation/planning stages and growth and maturity stages.

5.2.4. Multiple Regression Analysis

The multiple regression aims to determine the strength and direction of each independent variable's effect on institutional performance, while accounting for the other variables. The model is statistically robust ($F = 11.02$, $p < 0.001$), explaining 14.2% of the variance, a respectable result in behavioral and organizational research.

Table 4.
Multiple Regression values.

Variable	(β)	t-Value	p-Value	Interpretation
Technical Challenges	+0.577	5.81	<0.001	Remains the strongest predictor. Institutions that encounter greater technical challenges tend to achieve higher performance, reflecting a pattern of digital maturity.
Organizational Challenges	+0.161	2.26	0.024	Statistically significant but weaker effect. Suggests that governance and organizational structure play a secondary yet supportive role.
Recommendations	-0.153	-3.72	<0.001	Significant negative effect. Indicates that lower-performing institutions tend to place greater emphasis on the need for external recommendations.
Deployment Environment (On-Premises)	+0.120	4.21	<0.001	On-premises institutions report slightly higher performance compared to the reference category (likely cloud-based or hybrid).
Deployment Environment (Cloud-Based, Hybrid)	+0.045	1.34	0.181	Not statistically significant. No substantial performance difference from the reference category.
Constant	2.78	6.95	<0.001	Represents the baseline level of institutional performance in the absence of explanatory variables.

- Technical Challenges ($\beta = +0.577$, $p < 0.001$)

ISSN: 2576-8484
Vol. 10, No. 1: 60-76, 2026
DOI: 10.55214/2576-8484.v10i1.11493

This is the strongest predictor in the model. The standardized beta coefficient of +0.577 indicates that a one standard deviation increase in Technical Challenges correlates with a 0.577 standard deviation increase in Institutional Performance, representing a relatively large effect size.

This confirms the interpretation of Digital Maturity: technical challenges are not just a burden; they reflect significant institutional commitment. Institutions that invest in addressing these challenges (e.g., hiring specialists, procuring infrastructure, and developing custom models) generally achieve higher performance.

In this sense, technical challenges serve as a gateway to institutional excellence. Tackling them is not merely a cost but a strategic investment that directly enhances performance. It supports enabling infrastructure theory, which states that a robust infrastructure is essential for any successful digital transformation.

- Organizational Challenges ($\beta = +0.161$, $p = 0.024$)

There is a positive and statistically significant effect, but it is relatively modest compared to Technical Challenges. Even after considering other variables, improvements in governance and culture continue to contribute positively, albeit to a lesser extent.

This milder effect highlights that organizational challenges serve as enablers rather than main drivers. Good governance, by itself, does not ensure high performance, but it supports technologies and investments that enable effective and safe operation.

Organizational challenges are like the oil that reduces friction: they do not produce the output themselves, but they facilitate smoother, more sustainable performance. This aligns with Socio-Technical Systems Theory, which emphasizes the balance between technical and social/managerial aspects in achieving institutional efficiency.

- Recommendations ($\beta = -0.153$, $p < 0.001$)

Negative and statistically significant. This suggests that institutions that prioritize external recommendations tend to exhibit somewhat lower performance, even after accounting for technical and organizational challenges.

This reinforces the Institutional Saturation interpretation: high-performing institutions may find general recommendations less relevant or urgent in their advanced stage, as they may have already internalized or adapted policies locally and focus on more specialized optimizations rather than broad strategies. Recommendations play different roles depending on the institution's stage of adoption; they are more active in early planning phases and less so in later stages of refinement. This aligns with Knowledge Transfer Theory, which distinguishes between general and contextual knowledge, emphasizing how the importance of external guidance decreases as internal expertise develops.

- Deployment Environment (On-Premises: $\beta = +0.120$, $p < 0.001$ | Cloud: $\beta = +0.045$, $p = 0.181$)

The on-premises environment has a positive and significant influence on institutional performance compared to the reference environment (hybrid or unspecified), with a standardized beta of +0.120. This indicates that institutions operating locally tend to perform somewhat better, after accounting for other factors. The cloud environment does not have a statistically significant effect; its coefficient is small (+0.045), and the p-value (0.181) indicates non-significance.

This finding is intriguing as it supports the hypothesis that control over data and infrastructure is crucial for effective management. On-premises institutions benefit from complete control over their data, enabling more precise custom model training and avoiding some of the constraints commonly associated with cloud usage, such as variable costs, latency, and regulatory compliance.

Cloud deployments provide flexibility, but they can sometimes compromise customization or performance in sensitive applications. In situations, such as many in the Arab world, where data privacy and regulatory compliance are top priorities, local environments may offer a competitive advantage by delivering improved performance through tighter control.

This also aligns with the Resource-Based View in strategic management, which argues that control over strategic resources such as data and infrastructure offers an institutional advantage.

Technical challenges emerge as the strongest predictor of performance, both in correlation and in regression, indicating that institutions willing to engage deeply with technical complexity tend to achieve higher performance.

Organizational challenges also matter, but their effect is more modest; they function as supporting enablers rather than primary levers.

The negative correlation between recommendations and performance suggests that the value of external guidance is higher during early adoption stages and decreases as internal maturity advances.

On-premises environments appear to offer some performance benefits, likely due to greater control, privacy, and customization, particularly in regulatory or sensitive situations.

6. Results

- Technical Challenges Are the Strongest Predictor of Institutional Performance
  Both correlation and regression analyses show a statistically significant positive link between technical challenges and institutional performance. This suggests that institutions that invest in overcoming challenges such as data security, integration, and infrastructure tend to achieve better results.
  This supports the "Digital Maturity Paradox," where challenges reflect deep strategic adoption rather than just obstacles.
- Organizational challenges act as enablers rather than primary causes, while organizational factors such as governance and leadership have a statistically significant impact; their role is more supportive than causal. They provide the administrative support necessary for technology to function effectively.
- The Negative Correlation with Recommendations Reflects Maturity Levels
  Institutions with higher performance tend to rely less on external recommendations because they have already internalized best practices.
  In contrast, lower-performing institutions depend more on strategic guidance.
  This aligns with frameworks such as the Technology Adoption Lifecycle and Knowledge Transfer Theory, which highlight different needs at various stages of maturity.
- Deployment Environment Influences Performance Outcomes
  On-premises environments demonstrated a statistically significant performance advantage due to greater control over data and infrastructure.
  Cloud environments, while flexible, did not demonstrate a significant impact, likely due to challenges in customization, cost variability, and regulatory compliance in sensitive sectors.
- Near-Universal Agreement on the Importance of Future Recommendations
  Despite the weak negative correlation, the high mean and very low standard deviation suggest a widespread institutional consensus on the strategic necessity of policies and long-term governance frameworks for the adoption of LLMs.

7. Recommendations

- Invest in Flexible and Secure Technical Infrastructure
- Institutions should treat technical infrastructure, encompassing computing resources, cybersecurity, and integration capacity, as a strategic investment to enable the successful deployment of LLMs.
- Strengthen Governance and Executive Alignment with AI Strategy
  Develop clear institutional frameworks, train leadership on AI concepts, and activate dedicated AI governance units to oversee implementation and compliance.
- Align Deployment Environment with Institutional Readiness and Context
  Match infrastructure choices (cloud, on-premises, hybrid) with organizational needs,

ISSN: 2576-8484
Vol. 10, No. 1: 60-76, 2026
DOI: 10.55214/2576-8484.v10i1.11493

particularly considering that on-premises setups may offer competitive advantages in highly sensitive contexts.

- Reframe Technical Challenges as Opportunities for Excellence
- Treat technical barriers not as setbacks but as opportunities to build internal capabilities. Apply a Digital Enablement mindset to turn complexity into competitive differentiation.
- Tailor Recommendations to Institutional Maturity Levels
- Frame recommendations as foundational guidance for early adopters and offer refined, contextualized guidance for more advanced organizations.
- Launch comprehensive awareness and training programs Prioritize internal capacity-building through specialized training programs in AI governance, ethics, and domain-specific applications.
- Adopt a Digital Maturity Assessment Model for Ongoing Evaluation Use integrated assessment tools that measure technological, organizational, and performance maturity to track progress and guide strategic decisions.

8. Conclusion

This study provides a comprehensive analysis of the technical and organizational challenges associated with adopting large language models (LLMs) across cloud-based, on-premises, and hybrid environments within Arab enterprises. The findings reveal that each deployment model entails distinct trade-offs in terms of scalability, data security, operational control, and cost efficiency. Cloud environments offer agility and accessibility but raise concerns about data sovereignty. In contrast, on-premises solutions provide control at the expense of flexibility and scalability. Hybrid models appear to offer a balanced approach but demand advanced technical integration and governance. Organizational readiness, digital maturity, and stakeholder alignment are identified as critical success factors regardless of the deployment environment. The study contributes to the literature by contextualizing the adoption of LLMs within socio-technical frameworks and offers actionable recommendations to guide enterprise decision-makers in optimizing their AI strategies. Future research could further explore sector-specific implementations and the evolving role of open-source LLMs in enhancing enterprise AI adoption.

Institutional Review Board Statement:

This research was conducted in strict compliance with the 1964 Declaration of Helsinki and its subsequent amendments, as well as the guidelines of the researcher's institution. This research was not a medical study, nor did it involve human experimentation as outlined in the Declaration of Helsinki. All respondents in the study were over 18 years of age and voluntarily completed the research questionnaire. The information provided by the respondents was strictly used for this study and treated with the utmost confidentiality and anonymity.

Transparency:

The author confirms that the manuscript is an honest, accurate, and transparent account of the study; that no vital features of the study have been omitted; and that any discrepancies from the study as planned have been explained. This study followed all ethical practices during writing.

Acknowledgements:

I would first and foremost like to express my gratitude to all the participants who so kindly volunteered their views, which enabled this study to be conducted. I am also appreciative of our academic mentors and colleagues for their insightful comments and helpful critiques, which substantially improved this study. I also thank the early researchers whose theories formed the basis of our research methodology.

References

[1] F. Busch *et al.*, "Current applications and challenges in large language models for patient care: A systematic review," *Communications Medicine*, vol. 5, p. 26, 2025. https://doi.org/10.1038/s43856-024-00717-2

[2] Y. Zhang *et al.* "? Siren's song in the AI ocean: A survey on hallucination in large language models," *Computational Linguistics*, pp. 1-46, 2025.

[3] T. Chau, Y. Yan, and H. Liu, "A systematic review of AI ethics in education: Challenges, policy gaps, and future directions," *Journal of Global Information Technology Management*, vol. 28, no. 1, pp. 1–25, 2025.

[4] J. Swacha and M. Gracel, "Retrieval-augmented generation (RAG) chatbots for education: A survey of applications," *Applied Sciences*, vol. 15, no. 8, p. 4234, 2025.

[5] O. Gencoglu, A. Finn, R. Keshri, A. Salimoski, H. Honkanen, and J. Karhu, "Adoption of large language model AI tools in everyday clinical documentation: Mixed methods implementation study," *Journal of Medical Internet Research*, vol. 27, no. 1, p. e70789, 2025.

[6] TrueFoundry, "LLM in enterprise: A complete guide," 2024. https://www.truefoundry.com/blog/enterprise-in-llm

[7] Deepchecks, "How enterprises are deploying LLMs," 2024. https://www.deepchecks.com/how-enterprises-are-deploying-llms/

[8] Zemoso Labs, "Evaluating Llama and GPT: LLM adoption in enterprises," 2024. https://www.zemosolabs.com/insights/whitepaper/evaluating-llama-and-gpt-llm-adoption-in-enterprises

[9] Cloudflare, "What is an LLM (large language model)?," 2024. https://www.cloudflare.com/learning/ai/what-is-large-language-model/

[10] Al-Manara Consulting, "The general framework of scientific research," 2019. https://www.manaraa.com/post/3698

[11] BTS Academy, "How to prepare the general framework for scientific research," 2017. https://www.bts-academy.com/blog_det.php?page=137

[12] IBM, "What are large language models (LLMs)?," 2023. https://www.ibm.com/think/topics/large-language-models

[13] H. Naveed *et al.*, "A comprehensive overview of large language models," *arXiv preprint*, 2024. https://arxiv.org/abs/2307.06435

[14] Rethink Priorities, "Survey of developer experiences with large language models," 2024. https://rethinkpriorities.org/

[15] T. K. Dasaklis, F. Casino, S. Moschuris, and N. Kyriakou, "Large language models in human resource management: A systematic review," *SSRN Electronic Journal*, pp. 1-25, 2025.

[16] Wizr AI, "LLM evaluation: CIO's guide to enterprise AI models in 2025," 2024. https://wizrai.com/llm-evaluation

[17] J. M. Bodensohn, U. Brackmann, L. Vogel, A. Sanghi, and C. Binnig, "Unveiling challenges for LLMs in enterprise data engineering," *European Journal of Computer Science and Information Technology*, vol. 13, no. 31, pp. 65–75, 2024.

[18] Kong Research, *The state of enterprise LLM adoption: Investment trends, risks, and platform strategies in 2025*. San Francisco, CA: Kong Inc, 2025.

[19] Forbes, *Enterprise adoption of large language models accelerates amid security and regulatory concerns*. New York: Forbes Media, 2025.

[20] Nature Middle East, "Why we need Arabic language models. Nature Middle East," 2024. https://www.nature.com/nmiddleeast/

[21] H. Rhel and D. Roussinov, "Large language models and Arabic content: A review," *arXiv preprint arXiv:2505.08004*, 2025. https://doi.org/10.48550/arXiv.2505.08004

[22] O. Nacar *et al.*, "A culturally aligned benchmark for Arabic large language models," in *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing (pp. 1–15). Association for Computational Linguistics*, 2024.

[23] A. Rabehi, "Arabic LLMs in 2024: Benchmark & Gulf business applications. LinkedIn," 2024. https://www.linkedin.com/pulse/arabic-llms-2025-benchmark-gulf-business-applications-rabehi-phd-ml8kf

[24] F. Wei, D. Chen, Z. Liang, and S. Zhao, "Comparison study on the adsorption capacity of Rhodamine B, Congo Red, and Orange II on Fe-MOFs," *Nanomaterials*, vol. 8, no. 4, p. 248, 2018.

[25] T. Adenuga, N. Ayanbode, T. Ayobami, and F. C. Okolo, "Supporting AI in logistics optimization through data integration, real-time analytics, and autonomous systems," *International Journal of Scientific Research in Science, Engineering and Technology*, vol. 11, no. 3, pp. 511-546, 2024. https://doi.org/10.32628/IJSRSET241487