

Unlocking the Full Potential of AI in the Middle East

How Can AI Accelerate the Growth of the Region's Digital Economy?

Research by:

Melih Murat

Associate Research Director

Commissioned by:



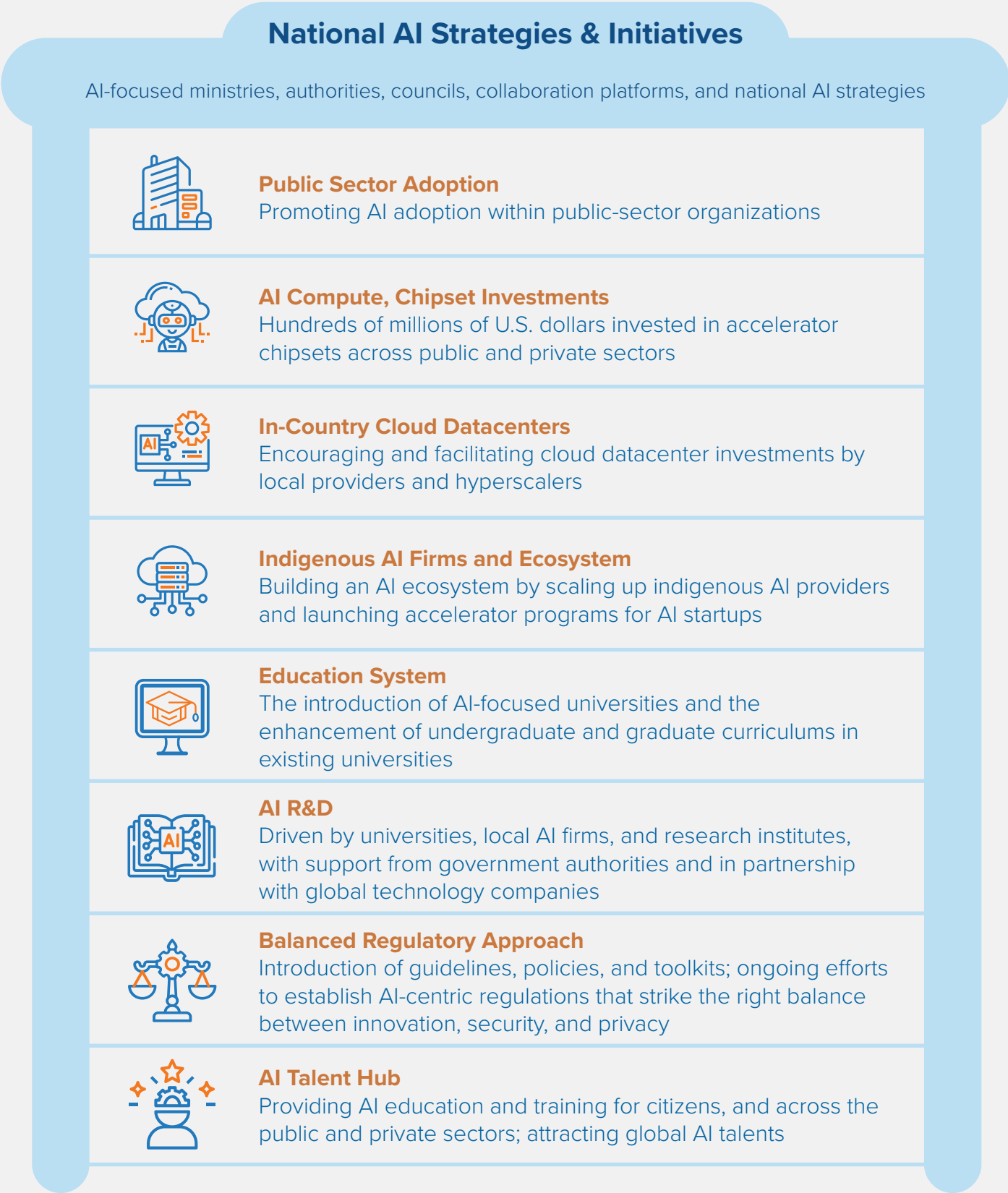


The Path to Achieving National AI Competitiveness and Independence

The global pursuit of a competitive AI ecosystem has become a priority for nations worldwide, and with good reason. Across the board, technology serves as the primary catalyst for economic growth, social advancement, and environmental sustainability. As countries vie to position themselves at the forefront of AI innovation, they recognize that mastery of this technology is essential to securing a prosperous and sustainable future.

Figure 1 below highlights key AI-related national initiatives across the Middle East. These initiatives have sparked a significant transformation across the region, evolving the countries from mere consumers of technology into active producers and innovators within the global technology landscape.

Figure 1: National Initiatives to Drive AI Competitiveness and Independence in the Middle East



Source: IDC and various online sources

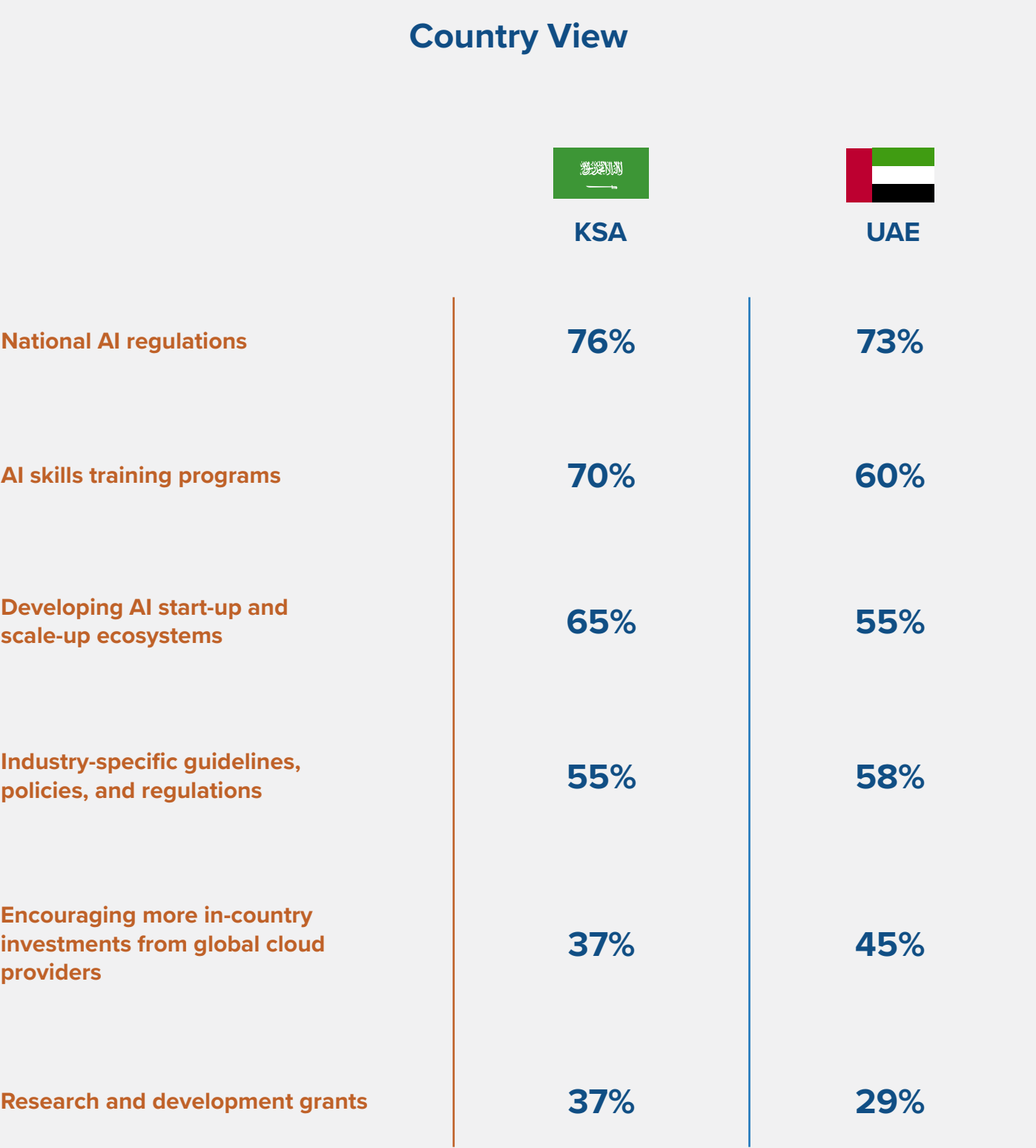
To gain insights into the trends surrounding private- and public-sector organizations’ perceptions and initiatives related to AI, AWS and IDC conducted an AI-focused survey in Saudi Arabia and the United Arab Emirates (UAE) involving 166 organizations across various industries, including hospitality and accommodation, financial services, government, media, entertainment, gaming, and retail and wholesale. This study primarily focuses on the UAE and Saudi Arabia and it serves as a valuable proxy for other technology-driven countries in the region. The key highlights of the study are as follows:



IDC also utilized other readily available market studies, where appropriate, to provide a comprehensive view of the trends.

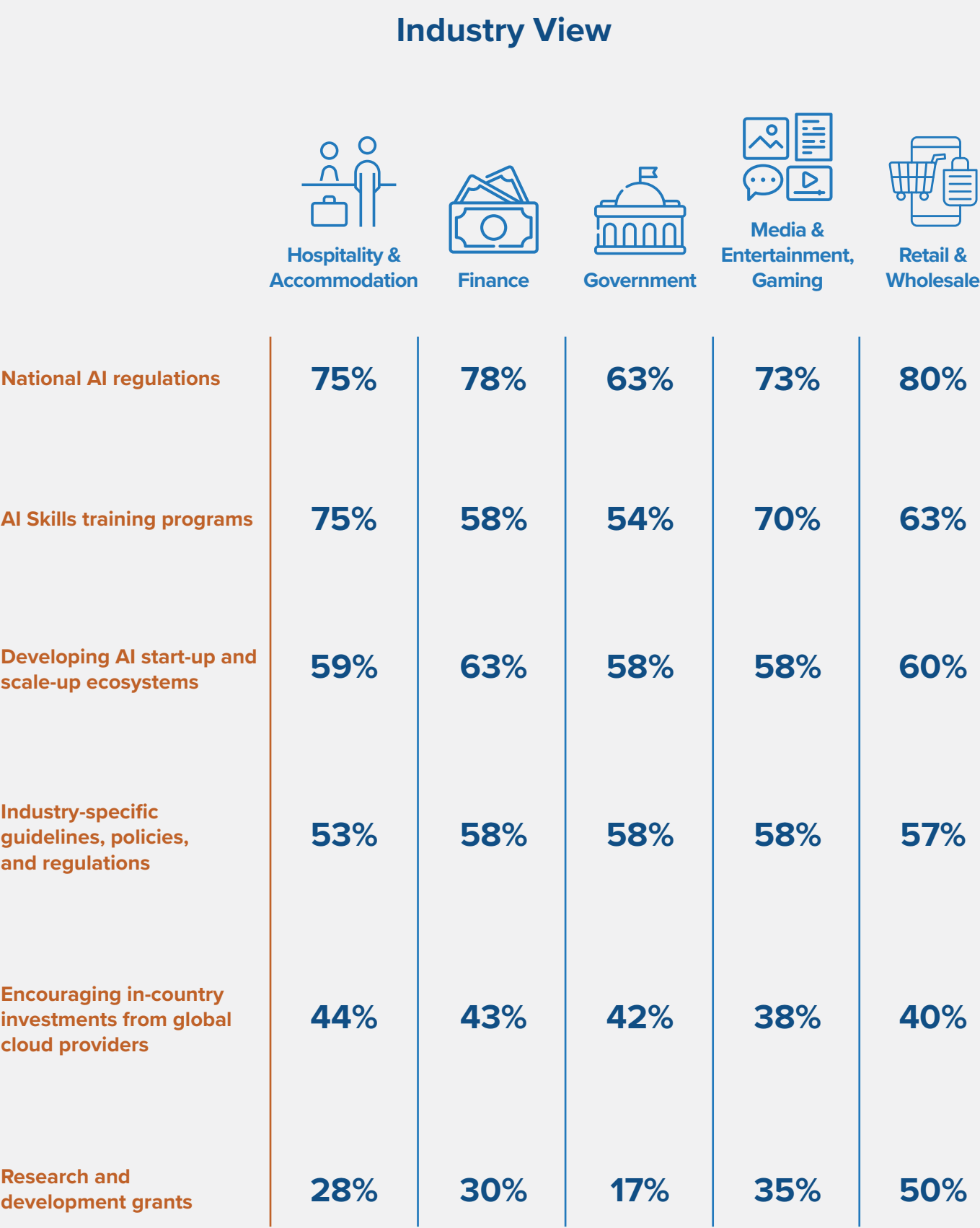
Figures 2a and 2b below offers insights into the government-led initiatives that are deemed important by public- and private-sector organizations in the UAE and Saudi Arabia.

Figure 2a: Awareness of Important and Impactful Government-Led AI Initiatives
Avoiding Overregulation of AI Market Strikes the Right Balance Between Compliance and Innovation
Q: What important and impactful government-led initiatives are you aware of that could drive your country’s competitiveness in the field of AI?



Source: AWS AI Readiness Survey for Saudi Arabia and UAE (Sample size: 166), August 2024

Figure 2b: Awareness of Important and Impactful Government-Led AI Initiatives
Avoiding Overregulation of AI Market Strikes the Right Balance Between Compliance and Innovation
 Q: What important and impactful government-led initiatives are you aware of that could drive your country’s competitiveness in the field of AI?



Source: AWS AI Readiness Survey for Saudi Arabia and UAE, August 2024 (N = 166)

Several important observations concerning some of the key trends are summarized below.



National AI Regulations

The majority of organizations emphasized the need for a regulatory framework to guide AI use case adoption across both critical and non-critical workloads, providing clarity for technology suppliers and IT buyers on how to address uncertainties surrounding mission-critical AI use cases and sensitive data, particularly considering regional and cultural nuances. They also stress the importance of avoiding overregulation, which could stifle innovation. Strict data residency requirements, for example, may limit organizations’ access to advanced public cloud-based AI technologies offered through global cloud datacenters, reducing a country’s global competitiveness. These challenges should be addressed through public consultations to ensure government authorities understand the expectations of both technology suppliers and buyers. Authorities should also engage with commercial public cloud providers to determine how and when these AI services can be offered through local datacenters, while also establishing a process to utilize global cloud datacenters until local services become available.



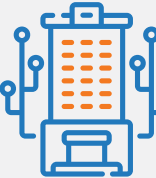
Skills Training Programs

Governments in the Middle East (e.g., Saudi Arabia, UAE, Qatar), in collaboration with universities, technology providers, and the private sector, have taken proactive steps to build national AI competencies among citizens, the public sector, and industries. 70% of organizations in the kingdom specifically highlight the importance of skills development initiatives to bridge the talent gap.



Industry-Specific Guidelines, Policies, and Regulations

While over half of the organizations in the UAE and in Saudi Arabia emphasized the importance of industry-specific regulations, it is crucial that industry authorities collaborate closely with national government bodies to ensure alignment between national regulations and those specific to the industry, particularly in areas directly or indirectly related to different type of AI use cases.



Developing AI Start-Up and Scale-Up Ecosystems

In recent years, the region has witnessed a surge in indigenous technology companies making a global impact. The survey underscores this as a vital initiative, contributing to a more comprehensive technology ecosystem. A notable example is the Falcon Foundation Model Family, developed by the Technology Innovation Institute in the UAE, which has emerged as a global success story and one of many locally developed technology solutions by regional companies.



Bolstering Local Datacenter’s Cloud Offering

It is important for regional authorities to encourage global cloud providers to offer AI platform and infrastructure services via local public cloud datacenters. Without cloud-based GPU and CPU compute resources, as well as data and AI platforms, AI innovation is impossible. Some countries, such as the UAE, Saudi Arabia, and Qatar, have made significant progress in this area by facilitating the development and launch of commercial public cloud datacenters (e.g., AWS), with more initiatives on the horizon across the region.

AI: Paving the Way for Enhanced Business Performance

88%

Of organizations in the UAE and Saudi Arabia that have already invested in AI, have reported improved business performance over the past two years. In contrast, companies still in the planning stages of AI adoption have fallen behind, showing considerably weaker results.

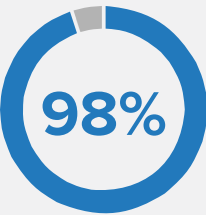
Source: AWS AI Readiness Survey for Saudi Arabia and UAE, August 2024 (N = 166)

State of AI Adoption and Investments in the Middle East

Investments across the AI value chain in the Middle East are strategic and comprehensive, targeting critical areas such as employee enablement, AI model and solution security, AI-enabled tools and applications, advanced analytics, AI at the edge, and data architecture modernization. Recognizing that AI tools only deliver value when employees are equipped to use them, there is also a strong focus on employee enablement and skills development.

Additionally, the analysis underscores the importance of foundational elements like data architecture modernization and a mature data management strategy, as these are crucial for successful AI integration. Overall, the investments are aligned with enhancing business outcomes, reflecting a balanced approach that combines technological advancement with the necessary human and structural support.

Strong Awareness of AI's Transformational Impact



of the surveyed organizations in the UAE and Saudi Arabia believe that AI will profoundly transform their businesses, driving them to prioritize AI-driven initiatives that demonstrably enhance business outcomes.

Source: AWS AI Readiness Survey for Saudi Arabia and UAE, August 2024 (N = 166)

Figure 3: Investment Priorities of Organizations Across the AI Value Chain
Organizations Embrace a Balanced Investment Approach by Blending Technology and Human Skills
Q: In which areas of your artificial intelligence value chain is your organization planning to invest the most in the next 12–18 months?

	 KSA	 UAE
Data literacy and AI skills enablement	69%	62%
AI solution/model security	60%	62%
AI-enabled tools/apps	54%	54%
Data architecture, engineering, and management	41%	42%
Advanced analytics and AI at the edge	41%	42%
AI-ready digital infrastructure and/or datacenters	29%	33%
Integrated AI use-case development	24%	21%
Change management	21%	13%
Responsible AI	19%	23%
AI platforms for model life-cycle management	12%	20%

Source: IDC Data and AI Survey for Saudi Arabia and UAE, August 2024 (N = 200)

AI Use Case Investment Plans

According to the study, finance, IT operations, and customer experience have emerged as the key business functions driving the majority of AI investments, reflecting their strategic importance in enhancing operational efficiency, innovation, and customer engagement. It is worth noting, however, that respondents indicated strong investment plans across all business functions.

In the rapidly evolving landscape of generative AI (GenAI, a subset of AI), business functions including corporate strategy, software development, marketing, and legal are currently at the forefront of adoption. As organizations set their sights on the future, customer service, facilities, and finance are expected to become the primary focus for GenAI investments. Notably, a substantial number of organizations are gearing up to integrate GenAI across a wide range of business functions.

The Acceleration of AI Adoption



of surveyed organizations in the UAE and Saudi Arabia, are currently investing in AI, while another 50% plan to invest in this technology. Conversely, 21% are not investing in AI and have no plans to do so.

Source: AWS AI Readiness Survey for Saudi Arabia and UAE, August 2024 (N = 166)

Figure 4: Generative AI Investment Plans by Business Function
Improved Decision Making, Employee Productivity, and Operational Efficiency are the Key Expectations from High-Priority Use Cases

Q: What is the state of GenAI investment at your organization within each of the following business functions?

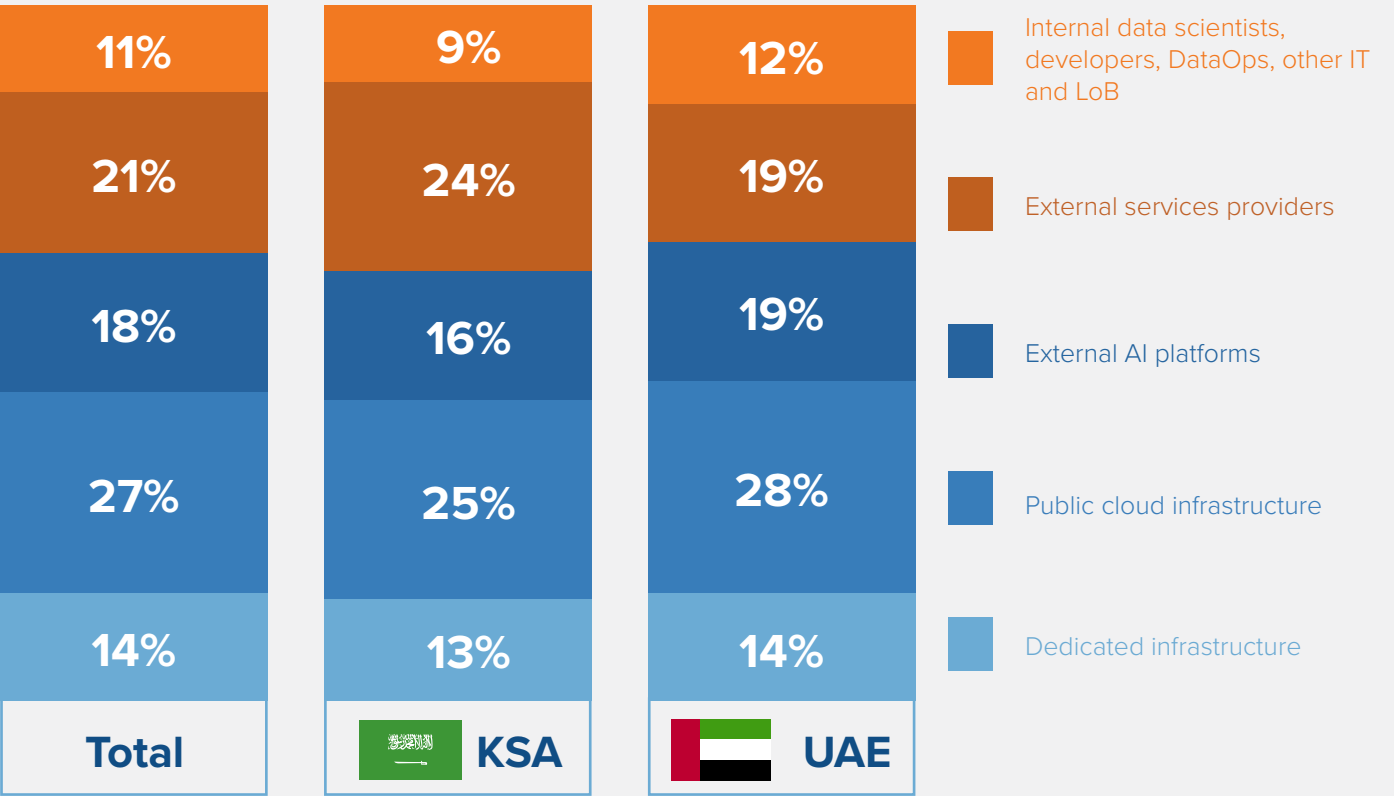
	Already Investing	Planning to Invest
 Corporate strategy	23%	41%
 Software development	11%	59%
 Legal	10%	55%
 Marketing	10%	44%
 IT operations	9%	60%
 Sales	7%	56%
 Engineering/R&D	7%	39%
 HR	6%	53%
 Procurement	6%	58%
 Security	5%	52%
 Supply chain	5%	63%
 Finance	4%	77%
 Facilities	4%	84%
 Customer service	3%	83%

Source: AWS AI Readiness Survey for Saudi Arabia and UAE, August 2024 (N = 166)

The growing adoption of generative AI is significantly influencing how organizations strategize their IT investments. Figure 5 below illustrates the distribution of GenAI spending across various areas.

Figure 5: Allocation of Generative AI Spending
Public Cloud-Enabled, AI-Ready Infrastructure Accounts for the Largest Share in the UAE and Saudi Arabia

Q: Thinking about your organization's new investment in generative AI-related projects in the next 18 months, what is your best estimate of how these investments will be allocated across the following areas?



Source: IDC Data and AI Survey for Saudi Arabia and UAE, August 2024 (N = 200)

Generative AI-related spending is heavily concentrated on IT infrastructure, underscoring the substantial resources required for model development, fine-tuning, and retrieval-augmented generation (RAG) initiatives. These efforts are primarily supported by public cloud compute resources, driven by the need for ongoing solution development and experimentation in the generative AI space.

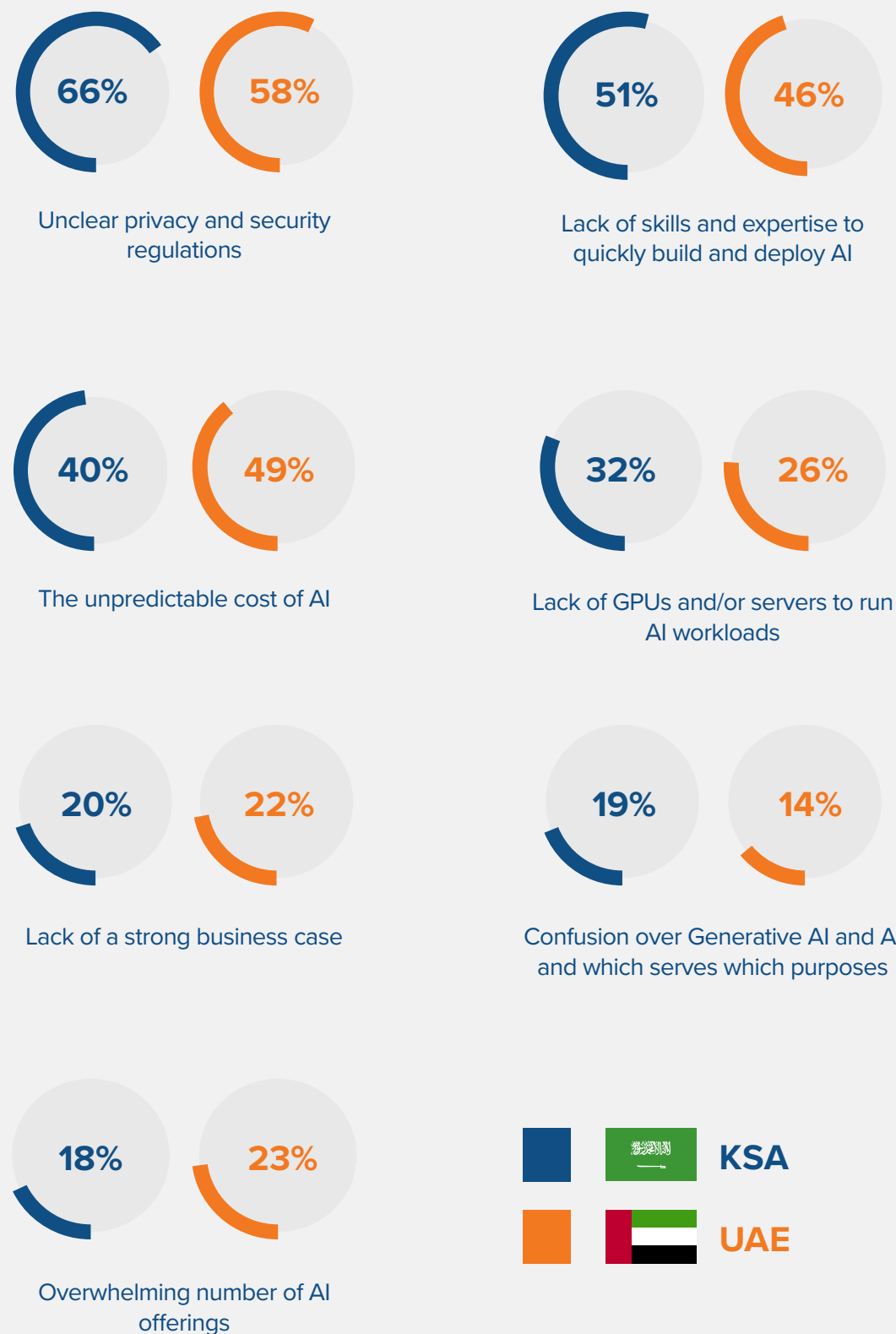
In addition to infrastructure, 18% of spending is directed towards external AI platforms to manage the entire model lifecycle, while another 21% is allocated to external service providers for the development, customization, deployment, and management of the complete GenAI technology stack. A further 11% of the budget is dedicated to empowering business and technology teams to spearhead GenAI initiatives.

Navigating the Challenges of AI-Enabled Digital Business

Figure 6: Limitations Hindering AI-Driven Digital Business Evolution

Technology Partners Play a Crucial Role in Resolving Most of These Challenges

Q: What limits your organization's willingness to accelerate AI adoption?



Source: IDC Data and AI Survey for Saudi Arabia and UAE, August 2024 (N = 200)



Lack of Clarity Across Privacy and Security Regulations: Protecting sensitive data is important, and AI systems, like any on-prem and public cloud services, should be secured properly. Organizations in the region often highlight that unclear privacy, security, and sovereignty requirements in existing regulations, along with uncertainty about upcoming regulations, hinder adoption of emerging technologies. It is also important to note that some organizations may lack the internal expertise to fully understand in terms of how they should secure their technology systems. In such cases, leveraging high-quality commercial public cloud datacenters, whether within or outside the country, becomes an attractive option, as they provide advanced security and privacy features that may surpass what organizations can achieve on their own.



Lack of Skills and Expertise: Building and deploying AI solutions requires specialized knowledge. A shortage of skilled professionals can slow down AI adoption and limit innovation. Employee enablement initiatives are a key method of tackling this challenge. Organizations can also leverage the consulting, professional, and managed services of AI technology and services providers to tackle this challenge.



Unpredictable Cost of AI: Without a proper governance mechanism and structured approach, like any other technology implementation initiative, AI projects can become expensive, due to unplanned costs related to infrastructure, development, and maintenance. Leveraging price calculation and cost management tools of technology partners, proactively planning and managing these expenses, and creating a list of AI use cases that will be implemented based on their potential business impact are important to justify further AI-related investments.



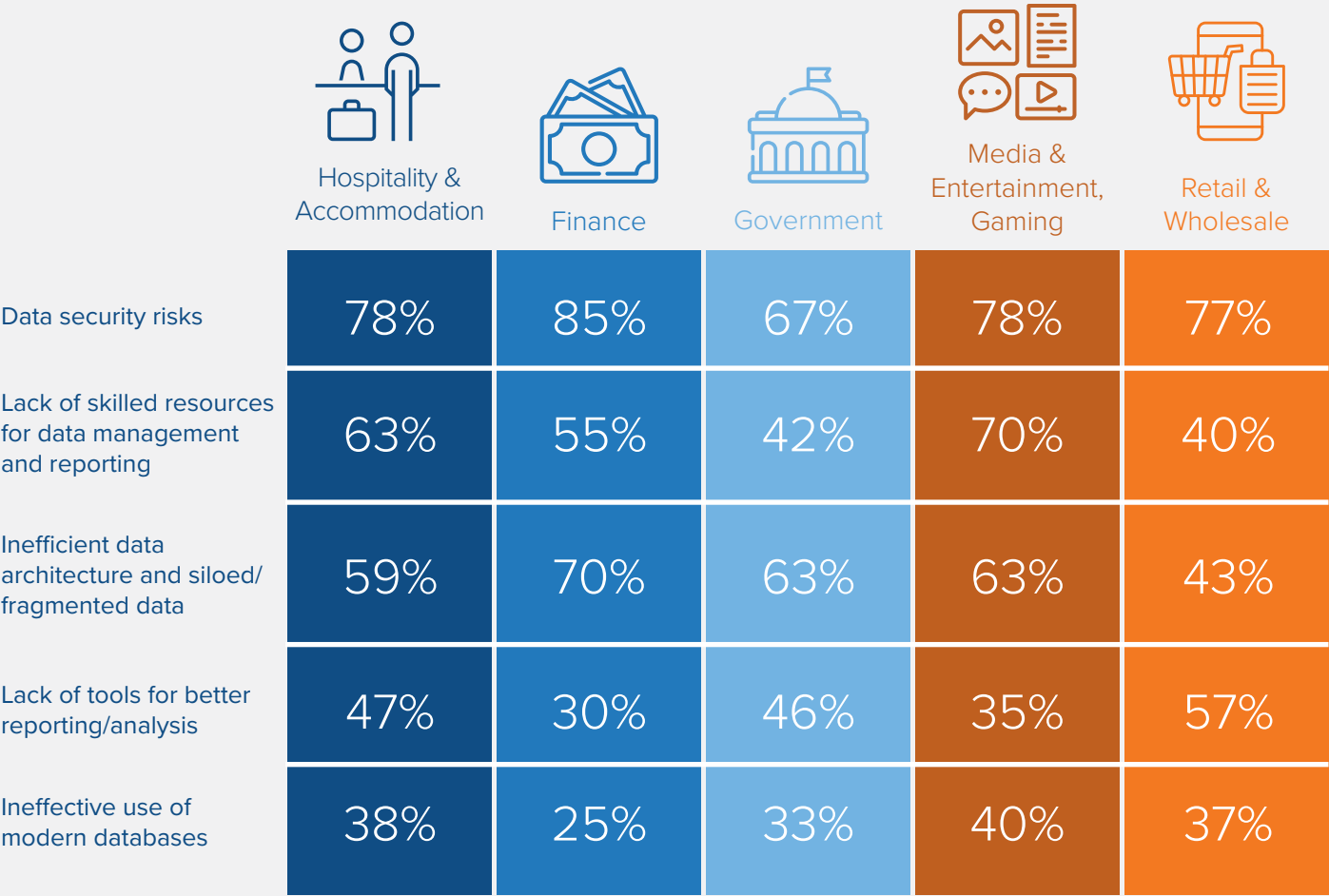
Lack of GPUs and/or Servers: AI workloads often require significant computational power. Limited access to GPUs or CPUs can bottleneck AI development and deployment, hindering progress. This is an area where authorities can take a proactive approach to the technology ecosystem in accessing these compute resources and providing support to IT buyers.



Lack of a Strong Business Case: Without a clear business case, securing buy-in for AI projects is challenging. Organizations need to articulate the tangible benefits of AI to justify investment. Working with competent technology services partners can also help in creating a list of use cases with a measurable business impact.

Data-related limitations are not addressed in Figure 7, as we chose to focus specifically on data challenges, given their foundational importance across all vertical markets. As evident from the survey results, the majority of respondents across various sectors highlighted key concerns, which included data security, a shortage of skilled resources for data management, and inefficient data architecture. These issues are critical pain points impacting the broader industry landscape.

Figure 7: Data-Related Challenges of Organizations
Overcoming Data Related Challenges Paves the Way for Faster AI Adoption and Time-to-Business Value
Q: What are your organization's top data-related challenges?



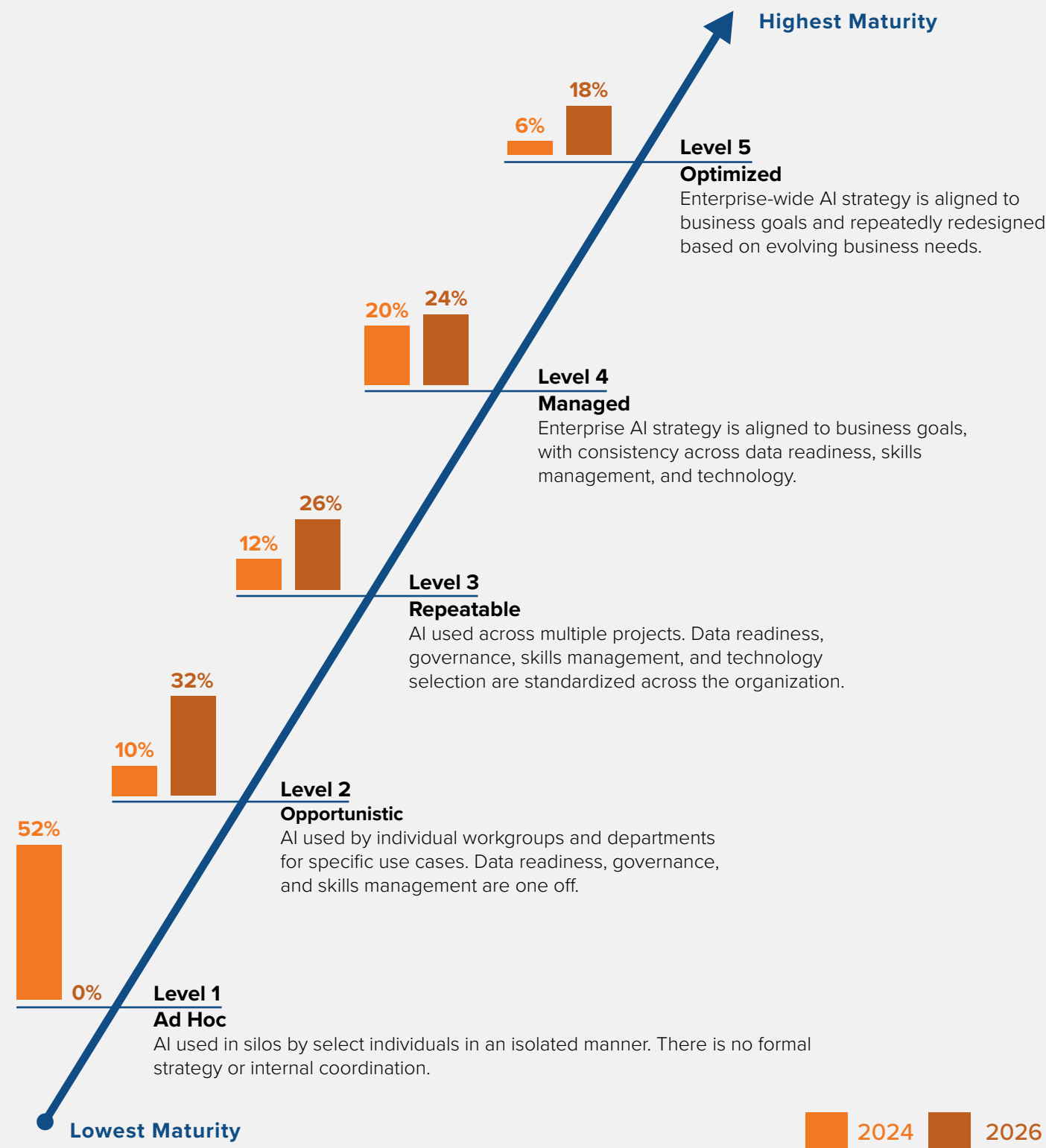
Source: AWS AI Readiness Survey for Saudi Arabia and UAE, August 2024 (N = 166)



Understanding the AI Maturity of Organizations in Middle East

To assess the AI maturity of organizations in the region, IDC used its AI maturity framework, which provides a structured, five-level approach to gauge respondents' current level of AI maturity. IDC mapped out where organizations in the UAE and in Saudi Arabia stand in their AI journey today, as well as their maturity goals for the next two years, offering a clear path for future AI growth. You can see the details of the framework and what each maturity level entails in Figure 8 below.

Figure 8: AI Maturity of Organizations: Current State vs. Target
Most organizations are at a lower maturity level but aspire to enhance their AI capabilities.
 Q: Which response below best represents your organization’s approach to artificial intelligence today and in two years’ time?



Source: IDC Data and AI Survey for Saudi Arabia and UAE, August 2024 (N = 200)

According to the analysis, most organizations currently evaluate themselves at lower maturity levels (levels 1 and 2), but many aim to progress to higher stages in the coming years. In the next section, we will explore what it takes to become a data- and AI-driven organization and successfully integrate AI.

Planning for Success: Foundational Pillars of AI-Led Digital Business

Below you can see the foundational elements that organizations need to focus on before embarking on their AI journey and what they will need to execute successfully with measurable business impact.

Figure 9: Road to Success in Artificial Intelligence
A Structured Approach to Driving AI Maturity and Impact

Business Readiness

Organizations Wide Data and AI Strategy Assess the organization’s AI readiness and develop a comprehensive AI strategy aligned with business objectives.	Executive Sponsorship and IT/ Business Alignment Foster collaboration across the entire C-suite and align business and technology teams to drive co-innovation.	Responsible AI Policy Create a framework that promotes the explainability, transparency, and integrity of AI models, ensuring their responsible and ethical use.	Employee Enablement Drive employee upskilling and reskilling initiatives to ensure the responsible development and deployment of AI-based solutions.
--	---	--	--

Technology Readiness

AI Ready, Cloud Native IT Infrastructure Utilize a hybrid cloud architecture to build, test, and deploy AI models that accommodate various computing, networking, and storage needs.	Data Architecture Modernization Adopt a modern data architecture to manage data efficiently and effectively, ensuring data readiness and eliminating data silos.	Privacy and Security Equip the organization to address the evolving threat landscape from both internal and external sources, while prioritizing data security and privacy.	AI Platform Deployment Implement a centralized AI platform to streamline data ingestion, preparation, and the entire lifecycle of AI model development, implementation, and management.
--	--	---	---

Innovation Culture Readiness

Use Case Selection Adopt a structured approach to use case identification by conducting design thinking workshops that foster collaboration both internally and externally.	Pilots and Use Case Prioritization Execute pilot projects to assess the potential impact of longlisted use cases, helping to shortlist those with the most significant benefits.	Use Case Implementation: From Test to Production Once the use cases are shortlisted, organizations can move from test to production and scale these use cases for organization-wide impact.	Lifecycle Management and Business Impact Analysis Regularly monitor the business impact of use cases and proactively manage their performance to ensure maximum impact on business.
---	--	---	---

AI-Centric Ecosystem Partnerships

with Technology Companies, Industry Stakeholders (e.g., Suppliers), Government Authorities, Universities, and Research Institutions to Foster Collaboration and Drive Co-Innovation.
 Ongoing Monitoring and Improvement of the AI Strategy and its Core Components

Source: IDC Data and AI Survey for Saudi Arabia and UAE, August 2024 (N = 200)

Below you can see the details of each of the three main readiness stages. Please note that some activities within each pillar are not sequential and can be executed concurrently.

Business Readiness

The first step in a successful AI transformation journey is to develop an organization-wide AI strategy that outlines key activities and milestones. Involving diverse business stakeholders from the beginning is essential to fostering an innovation-driven culture that aligns both business and technology teams. Cross-functional collaboration plays a critical role in breaking down silos and ensuring a unified approach to AI adoption. Furthermore, establishing a responsible AI framework during this phase is crucial, as it guarantees the safe, transparent, and unbiased deployment of AI technologies, fostering trust and accountability throughout the organization.



Technology Readiness

Many organizations in the region are focused on establishing a hybrid and multicloud digital infrastructure strategy to support a flexible, scalable, and agile digital business model. Equally important is the development of modern data architecture and strong data engineering capabilities. Without a cloud-enabled, AI-ready infrastructure and a robust data management strategy, organizations cannot begin their AI journey. Once infrastructure is in place and data readiness achieved, it is crucial to adopt an AI platform-led approach to manage the entire model lifecycle. Finally, organizations must remain proactive in addressing the evolving IT security risks while safeguarding data privacy.



Innovation Culture Readiness

For organizations to thrive in their AI journey and drive innovation, fostering a culture of experimentation is essential. This involves creating an environment where teams feel empowered to explore new AI solutions and use cases, take calculated risks, and learn from their experiences. In this culture, failure is viewed not as a setback but as a valuable learning opportunity. Teams are encouraged to quickly iterate on their ideas, refining their approaches based on real-world feedback. This agile way of working enables organizations to rapidly adapt and finetune their AI initiatives, ensuring they remain aligned with business objectives.

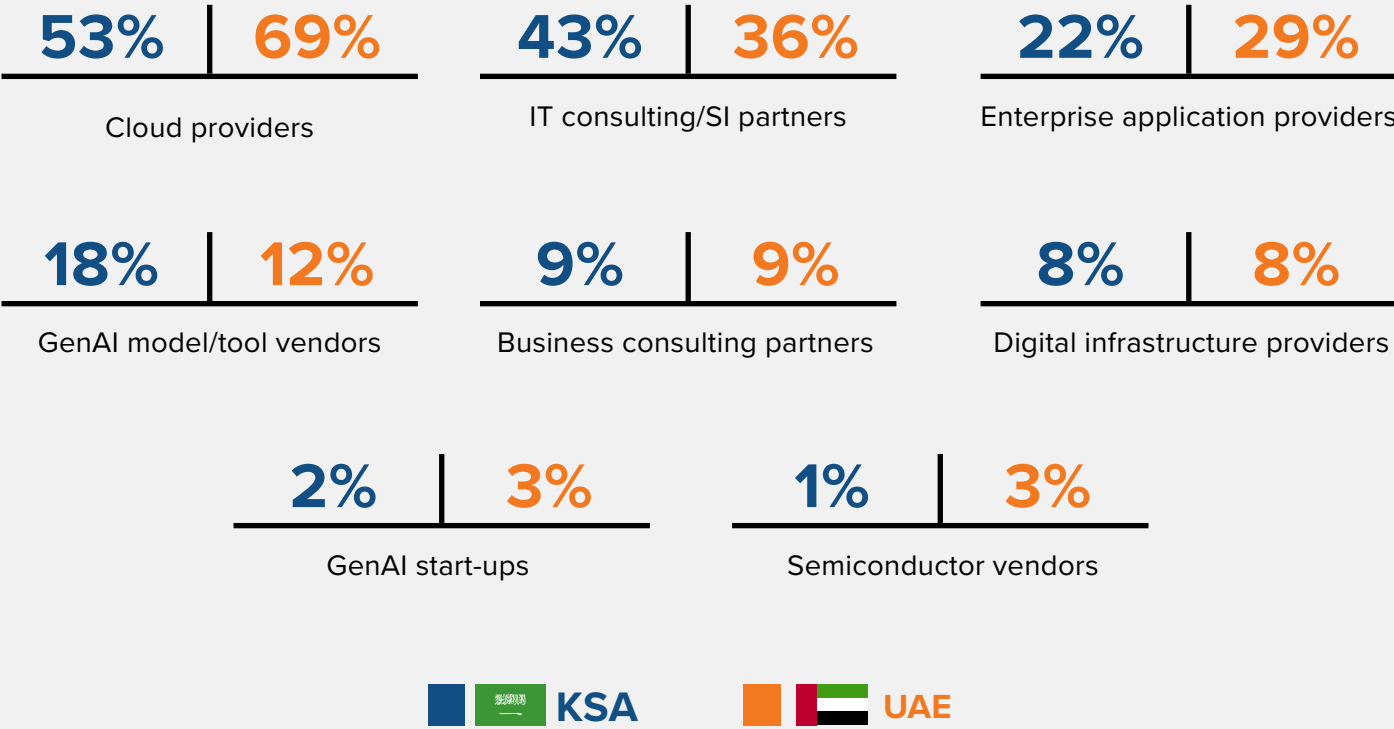


Significance of Building the Right Partnerships

An ecosystem-driven partnership approach is essential for the success of AI projects, as it brings together diverse expertise and resources from both technology and non-technology partners. Technology partners provide critical infrastructure, tools, and support, enabling organizations to implement advanced AI solutions effectively. Meanwhile, non-technology partners, including industry stakeholders, universities, and policy makers, bring collaboration opportunities that facilitate innovation. This multifaceted approach fosters a rich exchange of ideas and best practices, ultimately accelerating AI adoption and driving transformative outcomes across various sectors.

On the technology partnership front, cloud providers, followed by technology services companies, stand out as the most strategic partners of organizations in the UAE and Saudi Arabia when it comes to driving AI initiatives.

Figure 10: Most Strategic Generative AI Partners
Cloud Providers and Technology Services Companies Take Center Stage in the Middle East
Q: Which of the following vendor types will be your two most strategic GenAI technology partners in the next 12 months?



Source: IDC Data and AI Survey for Saudi Arabia and UAE, August 2024 (N = 200)

Partnerships with cloud providers offer AI-ready digital infrastructure, including servers, storage, and networking, as well as data and AI platform capabilities, such as Amazon SageMaker. Meanwhile, technology services companies assist organizations in achieving both business and technological readiness, guiding them throughout the entire AI journey — from use case selection and implementation to model lifecycle management.

Ensuring the Continued Success of AI Initiatives: An AI Platform-Led Approach

Both cloud providers and technology services companies should encourage their customers to choose and implement a data and AI platform, in order to ensure long-term success of AI initiatives. Such platforms integrate data management, analytics, and AI tools, providing seamless access to high-quality data, as well as model development, testing, tuning, and management capabilities. By leveraging these platforms, organizations can scale AI initiatives effectively, drive strategic insights, and realize substantial business value.

Figure 11: Evaluation of AI Platforms by Middle East Organizations
AI Platforms are Crucial for Succeeding in AI Adoption at Scale

Q: What is your organization’s plan for evaluating or using GenAI in the next 18 months?

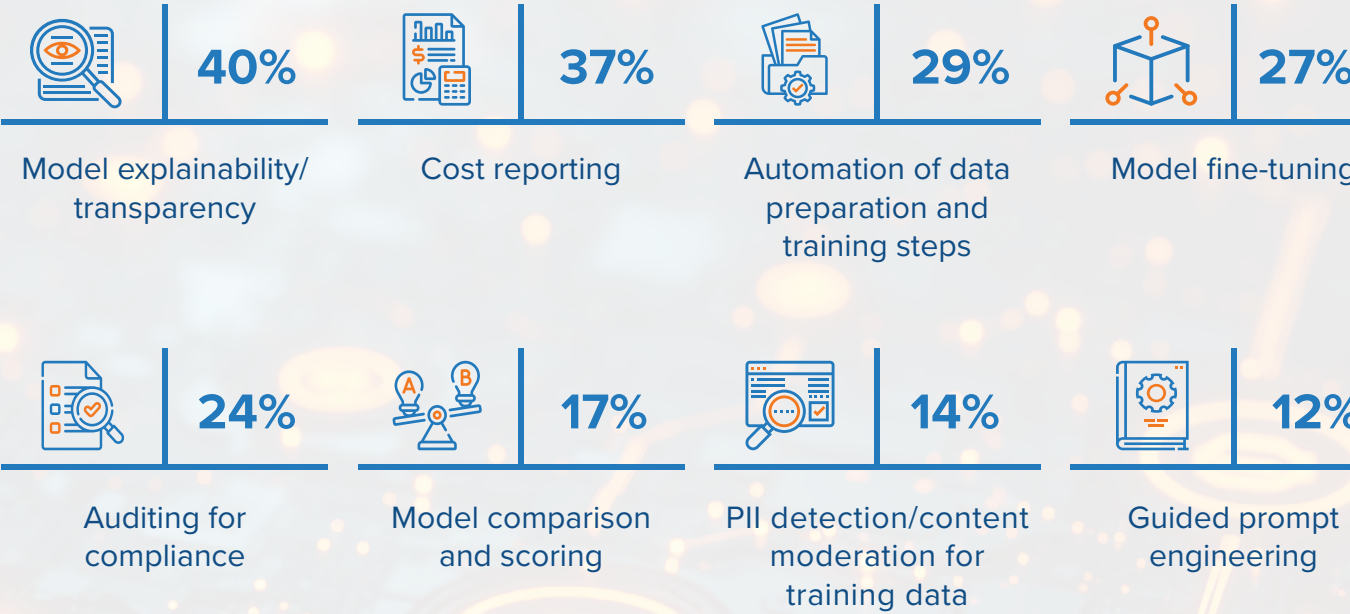
	 KSA	 UAE
Already selected a primary AI platform	32%	41%
Evaluating AI platforms and will select one in the next 3 to 6 months	39%	30%
Just starting to evaluate AI platforms and will select one in the next 12 months	27%	25%
Have not yet seriously evaluated any AI platforms	3%	3%

Source: IDC Data and AI Survey for Saudi Arabia and UAE, August 2024 (N = 200)

AI platforms play a critical role in ensuring model explainability and transparency, which is paramount for building trust and making informed decisions based on AI outputs. This is highlighted as the most important driver of investing in AI platforms. The second most important capability is that AI platforms facilitate cost reporting. This is also essential, as it provides a clear view of AI-related expenses, enabling better budgeting, resource allocation, and cost optimization. The third most important capability is the automation of data preparation and training, which is fundamental to drive data quality and efficiency in these processes.

Figure 12: Most Important Capabilities Expected in AI Platforms:
Model Transparency/Explainability, Cost Reporting, and Automated Data Preparation are the Top 3 Capabilities Expected in AI Platforms

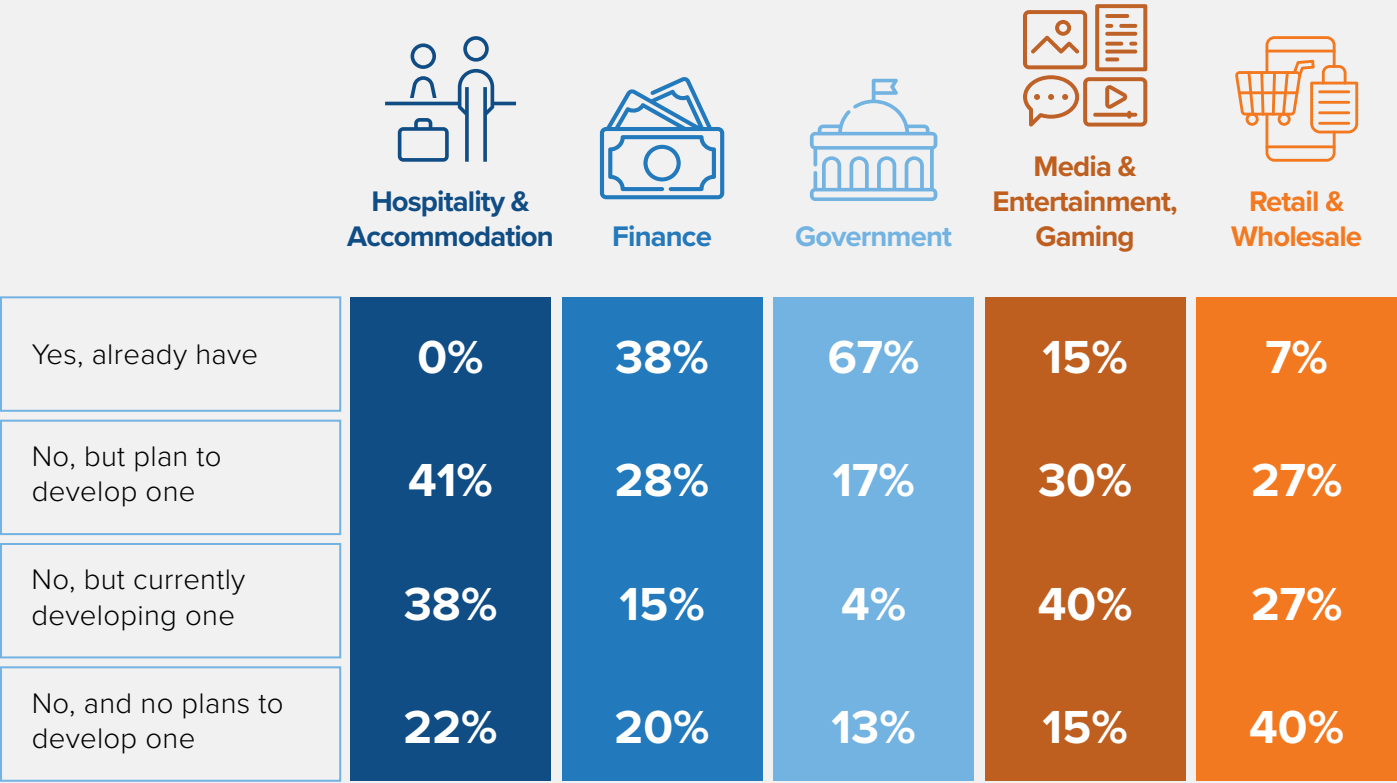
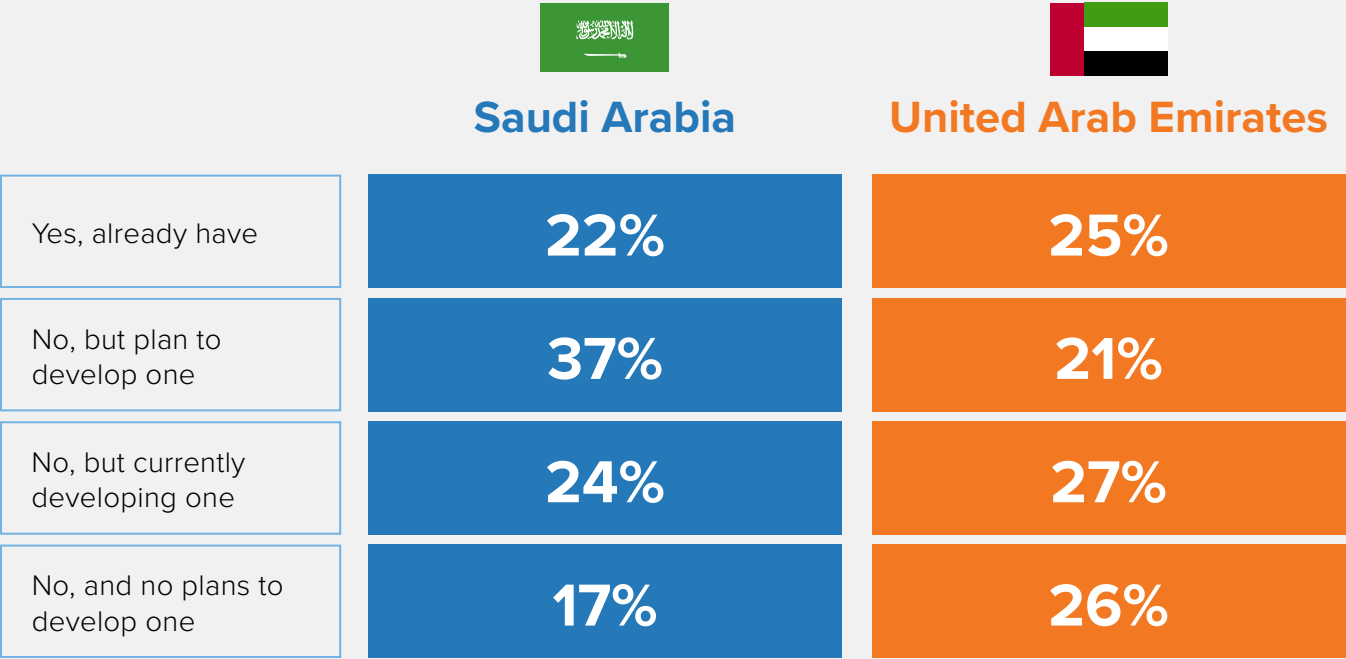
Q: Which of the following capabilities were, are, or will be the most important in your evaluation of AI platforms?



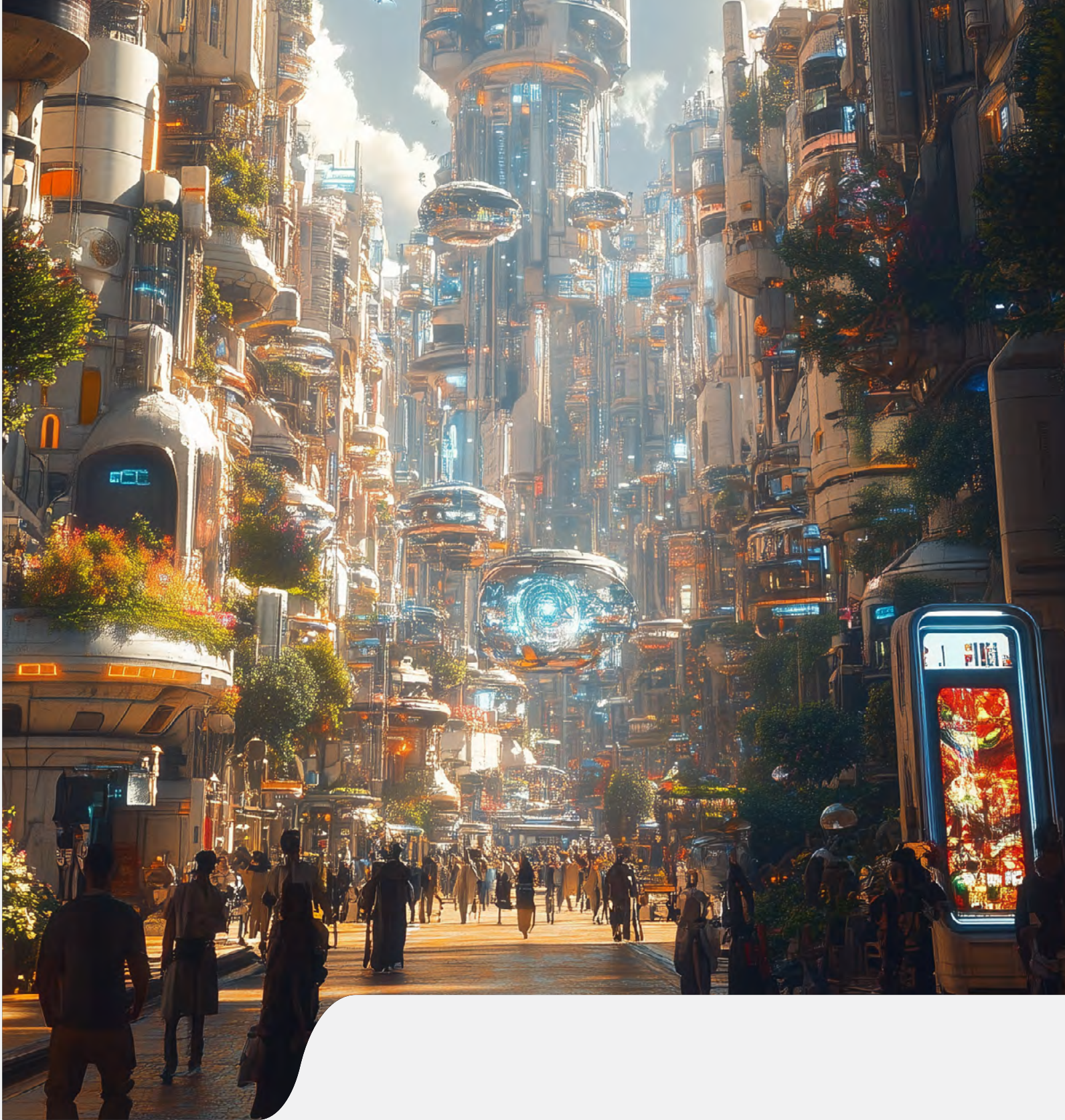
Source: IDC Data and AI Survey for Saudi Arabia and UAE, August 2024 (N = 200)

One of the crucial capabilities of these platforms is to support customers to ensure that the AI models developed by the customers adhere to regulatory standards and ethical guidelines, mitigating risks and safeguarding the organization’s reputation. Below you can see the current state of responsible AI initiatives. It is essential for organizations to establish a responsible AI policy before scaling AI adoption across their organizations.

Figure 13: Availability of a Responsible AI Policy
This is an Essential Step Before Scaling AI for Organization-Wide Use
Q: Does your organization currently have a responsible AI policy?



Source: AWS AI Readiness Survey for Saudi Arabia and UAE, August 2024 (N = 166)



Lessons Learned from
Successful Regional AI
Implementations

Camb.ai: UAE-Grown, World-Class AI Start-Up

Overview

Established in 2023 in the UAE, Camb.ai has rapidly emerged as a globally recognized start-up in AI-enabled content localization. Within a short time, the company’s innovative solutions have revolutionized multilingual movie dubbing and live content streaming (e.g., of sports events).

As a start-up, Camb.ai has greatly benefited from government programs and incentives, underscoring the UAE’s commitment to nurturing a vibrant, technology-driven ecosystem. During the discussions with IDC, the Camb.ai team raised an important concern regarding the availability of affordable and scalable GPU compute resources for start-ups and small businesses. According to Camb.ai CTO Akshat Prakash, AWS’s regional and global teams have been highly supportive in addressing this issue by including Camb.ai in number of start-up programs (e.g., cloud credits, technical support, and business mentorship), contributing to its business growth.

A Regional Partnership with Global Impact: Camb.ai and AWS

Camb.ai described AWS as a highly supportive partner in the development and delivery of its AI-based speech and translation services. AWS provided essential infrastructure for data storage and powerful training machines equipped with 8x NVIDIA H100 GPUs through AWS credits offered to Camb.ai. Additionally, AWS has supported Camb.ai’s go-to-market initiatives, helping the company reach as many potential clients as possible.

“AWS has been more than just a technology partner to us; it has been a trusted business ally and a business enabler. Its genuine commitment to our success is evident through numerous joint go-to-market initiatives. AWS’s support extends beyond the Middle East, contributing significantly to our global expansion.”

— Akshat Prakash, CTO, Camb.ai

AWS played an instrumental role in the development of Camb.ai’s generative AI-based speech models by providing essential compute resources (NVIDIA H100s) for training. In a notable recent development, Camb.ai unveiled its flagship MARS5 speech model at Dubai AI Retreat (May 2024). Developed from scratch, MARS5 has been open-sourced to foster further innovation. Camb.ai’s second foundational model, Boli, excels in contextual translation, ensuring correct grammar and appropriate colloquialisms for seamless translation between high resource and low-resource languages. Boli is currently proprietary but is expected to be open-sourced soon.

The Way Forward

Camb.ai’s journey has been marked by strategic partnerships, innovative AI solutions, and robust support from the UAE government and industry giants like AWS. Camb.ai plans to announce newer versions of its MARS model and expects to introduce more enhancements to its product lineup by leveraging its partnership with AWS.

The company is also working toward integrating its models with AWS Bedrock and other AWS marketplaces to streamline its go-to-market efforts. Camb.ai envisions becoming a leading AI player with Middle East origins, thereby showcasing the capabilities of UAE-based AI companies in the global market.

Technology Innovation Institute: Driving GenAI Innovation at a Global Scale

Overview

The Technology Innovation Institute (TII) is the Abu Dhabi government-funded global research center and applied research arm of the Advanced Technology Research Council (ATRC), which drives research and development strategy in the United Arab Emirates (UAE). One of TII’s research focus areas is generative AI (GenAI), and the institution has been developing open-source large language models to drive innovation in this area. To achieve this, TII has already built a diverse team of technology experts from 20 different nationalities and fostered strong partnerships within the global technology ecosystem. This diverse and collaborative environment enables groundbreaking research and development, contributing to the UAE’s growing reputation as a global technology innovation hub.

TII developed the Falcon Model Family, a series of state-of-the-art large language models (LLMs) that have garnered global recognition for their performance and versatility, using the AWS infrastructure. In 2023, TII launched Falcon 7B, 40B, and 180B, the latter two topped the global HuggingFace leaderboard rankings at the time of the release. In 2024, TII introduced Falcon 2, including the Falcon 2 11B and the Falcon 2 11B vision-language model (VLM), and Falcon-Mamba 7B, both topped again the rankings on the global leaderboards.

Fueling Generative AI Innovation Worldwide: Amazon Web Services & TII Collaboration

TII’s partnership with AWS has been instrumental in the development and deployment of the Falcon model family. The collaboration between AWS and TII began towards the end of 2021, driven by AWS’s advanced and comprehensive infrastructure-as-a-service (IaaS) offering. The IaaS platform was distinguished by its ability to meet the architectural requirements of TII for building and training the Falcon model family.

“AWS provided a robust infrastructure, eliminating the need to address infrastructure-related challenges and allowing us to focus on model development and testing. Moreover, our relationship with AWS extends beyond infrastructure, encompassing R&D projects and commercial initiatives through TII’s sister companies.”

— Dr. Hakim Hacid, Acting Chief Researcher, AI Cross-Center Unit at TII

TII leveraged AWS technologies, including SageMaker for model training, S3 for storage, and EC2 instances for various computational needs. The Falcon model family was trained on AWS SageMaker using the following compute capacity:

Falcon-7B	was trained on 384 A100 40GB GPUs
Falcon-40B	was trained on 384 A100 40GB GPUs
Falcon 180B	was trained on 3.5 trillion tokens on up to 4,096 A100 GPUs
Falcon2-11B	was trained on 1,024 A100 40GB GPUs
Falcon2-11B-VLM	was trained on 16 A100 80GB GPUs
Falcon-Mamba 7B	was trained on 256 H100 80GB GPUs

Additionally, the Falcon model family is available through AWS’s global cloud regions on SageMaker JumpStart, a platform that provides pretrained, open-source models for a wide range of problems. This makes the Falcon model family accessible to AWS customers worldwide.

The Way Forward

Looking ahead, TII plans to continue its innovation trajectory with several key initiatives:



Mixtures of experts
Developing advanced models utilizing the mixtures of experts technique to enhance performance and scalability.



Reinforcing multimodality
Expanding the capabilities of Falcon models to integrate more modalities beyond text and images.



Enhancing reasoning capabilities
Exploring ways to embed sophisticated reasoning abilities within the models.



Security enhancements
Incorporating improved security measures to address risks such as prompt injection.

TII is committed to leveraging its strategic partnership with AWS, ensuring continuous access to essential infrastructure and support for ongoing and future projects. AWS and TII are continuing discussions to explore joint go-to-market initiatives to drive global adoption of the Falcon model family. This collaboration aims to replicate successful projects and scale them to meet global demand, reinforcing the UAE’s position as a leader in AI and technological innovation. Through these efforts, TII not only aims to advance AI technology, but also to create a historical impact on the technology market by demonstrating the UAE’s capacity to contribute significantly to global tech advancements.

Almosafer: An AI Driven Omni-Channel Travel Services Provider

Overview

Almosafer, Saudi Arabia’s leading travel company (part of Seera Group), serves travelers through seamless omni-channel consumer experiences via web platforms, mobile applications, WhatsApp, retail branches and a contact center. With services that include flight bookings, hotel reservations, holiday packages, car rentals, travel insurance, activities, and more, Almosafer’s consumer business is now the leading travel brand in KSA and has become a well-recognized name across the Middle Eastern travel industry.

As a digital first company, Almosafer has been leveraging its partnership with AWS since 2019 to transform and elevate its core travel platform and other critical workloads. The collaboration initially focused on re-architecting Almosafer’s technology stack, transitioning legacy applications to Amazon Elastic Kubernetes Service (EKS), and rapidly developing new products and services at scale. This strategic move has significantly enhanced Almosafer’s business agility and responsiveness to customer needs. Furthermore, it has laid a solid foundation for the accelerated integration of AI into various internal processes and customer services, driving innovation, improving customer experience, and boosting operational efficiency.

Almosafer Leveraging AWS’ AI Platform for Accelerated AI Infusion

After establishing the foundational cloud infrastructure with AWS and modernizing legacy applications, the next step for Almosafer was to infuse AI into these applications and services to enhance and personalize the customer experience. To achieve this, Almosafer implemented several organizational initiatives, which are outlined as follows:



Created a structure for enabling the collaboration of cross-functional IT and business teams based on a centralized data and AI strategy



Secured C-level support for prioritizing and implementing AI use cases



Built an in-house data science team to drive AI innovation based on business needs



Identified success criteria based on measurable key performance indicators (KPIs)

On the technology front, Almosafer has been leveraging the AWS platform to manage the entire data lifecycle and to develop, deploy, and manage AI solutions. On the data management front, Almosafer utilized Amazon EMR for cloud-based big data management, AWS Glue for serverless data integration, AWS Athena as an interactive query service to analyze data on Amazon S3, and AWS Lake Formation for centrally governing, securing, and sharing data for analytics and machine learning (ML). In terms of AI initiatives, Amazon SageMaker was selected to manage the entire AI use case lifecycle.

“AWS excels in data integration, allowing us to easily consolidate and access information across internal and external data sources. AWS SageMaker’s comprehensive AI platform equips our data science team with the requisite tools to drive customer-centric innovation by leveraging data from various sources in a secure and trusted manner.”

— Bashar Khatib, Head of Data Engineering at Almosafer

The Way Forward

Based on the successful impact of its partnership with AWS, Almosafer has decided to expand this collaboration by leveraging Amazon Bedrock. Almosafer aims to use the Amazon Bedrock platform to manage the lifecycle of its GenAI projects in a more responsible, agile, secure, and reliable manner. The Almosafer team has noted that their initial interactions with Amazon Bedrock demonstrated significant potential for transforming both front-end, customer-facing and back-end operational processes through an AI platform-led approach. AWS has already conducted training sessions on the Amazon Bedrock platform for the Almosafer team and provided credits to enable their initial experimentation with the platform. This expanded partnership promises to drive continued innovation and operational excellence at Almosafer, setting the stage for further customer experience and business efficiency advancements.

Key Takeaways for Government Authorities



Lead by Example in Driving AI Adoption: Promoting AI adoption within the public sector, in alignment with national AI strategies, is crucial for ensuring successful program execution. This will also motivate private-sector organizations to pursue AI initiatives more decisively. Public-sector entities can effectively communicate the benefits and value of a structured and responsible approach to AI implementation.



Enable Access to AI-Ready Cloud Infrastructure and Platform Capabilities: AI-ready infrastructure and platforms are critical for driving innovation in AI. Some governments in the region have successfully encouraged investment of local and global cloud datacenter services providers. However, it is important to acknowledge that certain digital AI infrastructure and platform capabilities may not yet be available in domestic cloud datacenters right away. This is understandable, given the unprecedented pace of technological evolution, as the newest cloud services are first introduced in global hyperscaler datacenters, later arriving in domestic datacenters. Authorities can provide clear regulatory guidelines for accessing cloud datacenters, both domestically and internationally, to ensure that essential infrastructure and platform capabilities remain accessible to organizations. These authorities can also work closely with the local and global public cloud providers to build the business case for those capabilities to be made locally available



Pursue AI-Related Regulatory Initiatives with an Innovation-First Approach: This study highlights the importance of addressing gaps in organizational understanding of certain regulations, particularly those concerning privacy and security, while providing guidance for responsible AI adoption. When establishing or enhancing regulations, governments should proactively engage with technology suppliers and buyers, gathering their insights and recommendations early in the process. This ensures that regulations foster, rather than hinder, innovation, such as by enabling controlled access to the latest global AI infrastructure and platform capabilities that may not be available within the country.



Foster International Collaborations for Global Competitiveness: Governments can enhance global AI competitiveness by participating actively in international AI forums and treaties and by establishing bilateral and multilateral partnerships to share knowledge and resources. Collaborating with other nations on initiatives like AI research and regulatory alignment will help boost competitiveness on the global stage.



Facilitate Industry Ecosystem Partnerships for Complex National Initiatives: Adopting an ecosystem-driven approach — bringing together policymakers, technology suppliers, buyers, research institutions, universities, and startups — can help address national technological challenges. Government authorities play a pivotal role in fostering industry collaboration by facilitating partnerships that develop AI solutions for national issues and encourage corporate investment in AI technologies and R&D.



Continue Focusing on Human Capital Development: The rapid evolution of AI technologies necessitates continuous training and development. Integrating AI and digital skills into national education curricula at all levels, as well as establishing AI-specific degrees, certification programs, and apprenticeships, are efforts that have been underway for several years. Government authorities should continue pursuing nationwide programs to reskill and upskill workers in AI-related fields. Additionally, creating policies that attract top-tier global AI researchers, engineers, and entrepreneurs can help close the talent gap and accelerate local innovation.

Key Takeaways for Technology Buyers



Develop an Organization-Wide AI Strategy: Cross functional teams should collaborate to craft a comprehensive AI strategy that aligns with business objectives. This strategy should outline key milestones and activities while setting a clear path for AI adoption.



Establish Organizational Guardrails & Policies: Establish clear organizational policies and governance frameworks to guide AI adoption. These guardrails should focus on ethical AI use, data privacy, and security, ensuring compliance with local and international regulations.



Cultivate Internal Skills and Foster an Innovation-Driven Culture: To ensure successful AI implementation, IT buyers should invest in upskilling and reskilling initiatives that build internal AI expertise. Fostering a culture of innovation across teams will not only encourage collaboration but also drive creative solutions that leverage AI technologies.



Design a Cloud-Native Digital Infrastructure: IT buyers must prioritize building a cloud-native infrastructure that supports scalability, flexibility, and agility. By adopting hybrid cloud strategies, organizations can enhance their ability to innovate while meeting diverse business needs.



Build a Modern Data Architecture: A modern data architecture is foundational for AI readiness. IT buyers should focus on creating a scalable, robust framework that eliminates data silos and supports real-time data access. This architecture should allow for efficient data ingestion, preparation, and management.



Implement an AI Platform for Data and Model Management: To streamline the development, deployment, and maintenance of AI models, IT buyers should adopt an AI platform-driven approach. This approach ensures end-to-end model lifecycle management, from data ingestion to model retraining. It offers greater control, automation, and scalability, optimizing the use of AI and improving business outcomes.



Embrace an Agile, Iterative Approach to Use Case Selection and Execution: IT buyers should adopt an agile and iterative approach when selecting and implementing AI use cases. This method allows for continuous testing, learning, and refinement, ensuring that the most impactful use cases are prioritized.



Create a Technology Ecosystem for Accelerated Time to Value: Building a robust technology ecosystem is essential for accelerating AI deployment and driving business value. IT buyers should collaborate with cloud providers, technology services companies, software and hardware vendors, and other technology players to create an interconnected framework that supports rapid experimentation and innovation. It is also important to connect with non-technology stakeholders (e.g., government entities, universities, industry peers) to develop and implement industry-specific AI solutions.

About AWS

Since 2006, Amazon Web Services has been the world's most comprehensive and broadly adopted cloud. AWS has been continually expanding its services to support virtually any workload, and it now has more than 240 fully featured services for compute, storage, databases, networking, analytics, machine learning and artificial intelligence (AI), Internet of Things (IoT), mobile, security, hybrid, media, and application development, deployment, and management from 108 Availability Zones within 34 geographic regions, with announced plans for 18 more Availability Zones and six more AWS Regions in Mexico, New Zealand, the Kingdom of Saudi Arabia, Taiwan, Thailand, and the AWS European Sovereign Cloud. Millions of customers—including the fastest-growing startups, largest enterprises, and leading government agencies—trust AWS to power their infrastructure, become more agile, and lower costs.



To learn more
aws.amazon.com

About IDC

International Data Corporation (IDC) is the premier global provider of market intelligence, advisory services, and events for the information technology, telecommunications, and consumer technology markets. IDC helps IT professionals, business executives, and the investment community make fact-based decisions on technology purchases and business strategy. More than 1,300 IDC analysts provide global, regional, and local expertise on technology and industry opportunities and trends in over 110 countries worldwide. For over 50 years, IDC has provided strategic insights to help our clients achieve their key business objectives. IDC is a subsidiary of IDG, the world's leading technology media, research, and events company.

IDC Custom Solution

This publication was produced by IDC Custom Solutions. The opinion, analysis, and research results presented herein are drawn from more detailed research and analysis independently conducted and published by IDC, unless specific vendor sponsorship is noted. IDC Custom Solutions makes IDC content available in a wide range of formats for distribution by various companies. A license to distribute IDC content does not imply endorsement of or opinion about the licensee.



IDC Middle East/Africa

Level 15, Thuraya Tower 1 Dubai Media City
P.O. Box 500615
Dubai, United Arab Emirates
+971.4.3912741
Twitter: @IDC
idc-community.com
www.idc.com

Copyright Notice

Permissions: External Publication of IDC Information and Data

Any IDC information that is to be used in advertising, press releases, or promotional materials requires prior written approval from the appropriate IDC Vice President or Country Manager. A draft of the proposed document should accompany any such request. IDC reserves the right to deny approval of external usage for any reason.

Copyright 2024 IDC. Reproduction is forbidden unless authorized. All rights reserved.