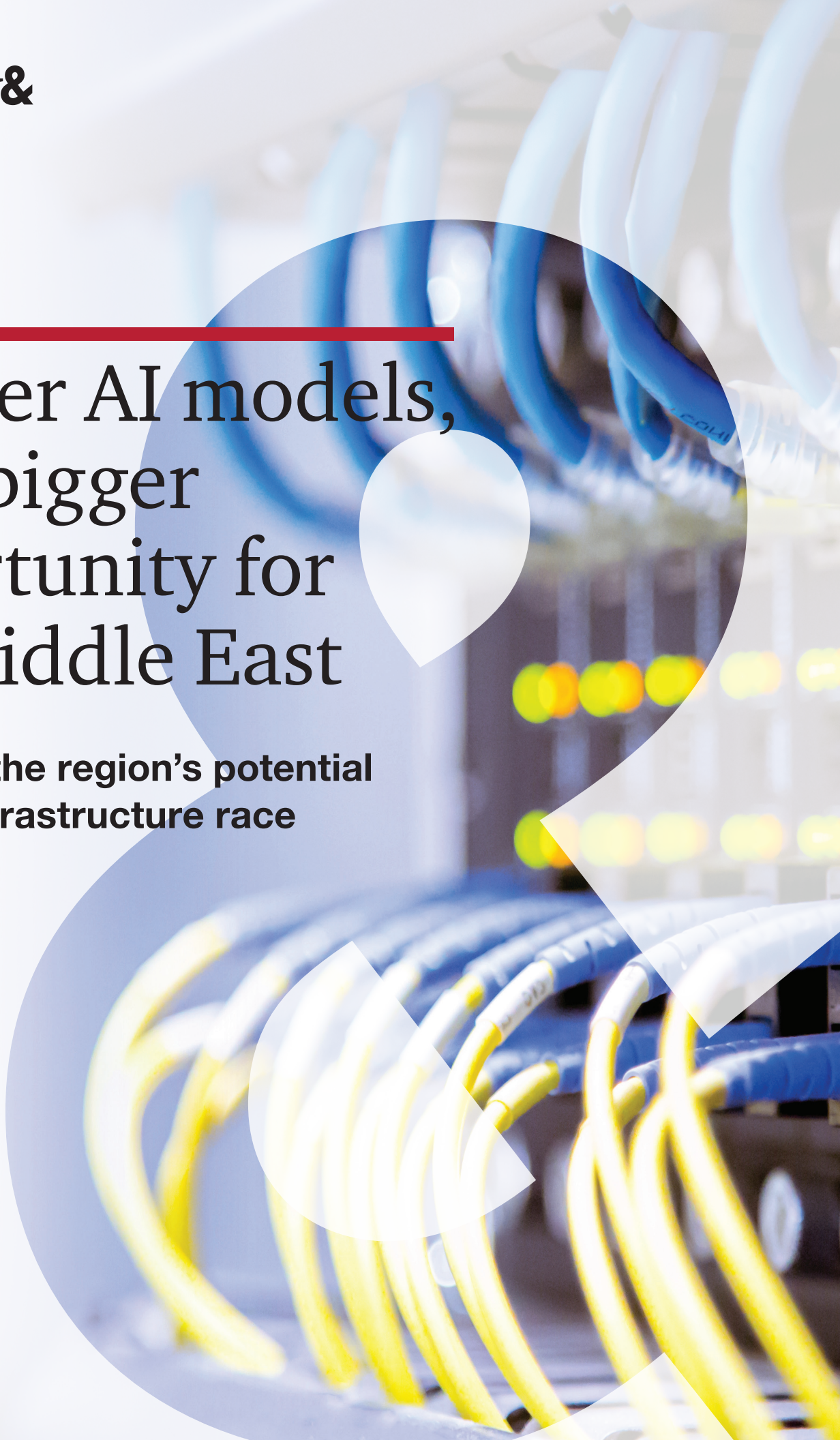
strategy&
Part of the PwC network



Smaller AI models, even bigger opportunity for the Middle East

Unlocking the region's potential in the AI infrastructure race

Contacts

Dubai
Imad Atwi
Partner
+971-4-436-3000
imad.atwi
@strategyand.pwc.com

Prateek Chauhan
Principal
+971-4-436-3000
prateek.chauhan
@strategyand.pwc.com

Riyadh
Diana Dib
Partner
+966-11-249-7781
diana.dib
@strategyand.pwc.com

Kirolous Zikry
Principal
+966-11-249-7781
kirolous.zikry
@strategyand.pwc.com

About the authors

Diana Dib is a partner with Strategy& Middle East, part of the PwC network. Based in Riyadh, she is a member of the technology, media, telecommunications, and digital practice in the Middle East. She has over 18 years of management consulting experience across Europe and the Middle East. She specializes in the tech and digital sector, advising brownfield and greenfield technology and artificial intelligence (AI) champions and ventures, national investment authorities, and sector developers. Her expertise spans AI business models and go-to-market strategies, digital infrastructure, and data center investments and plays.

Imad Atwi is a partner with Strategy& Middle East and a member of the technology, media, telecommunications, and digital practice in the Middle East. He advises sovereign entities, national champions, and telecom operators on digital infrastructure, tech and digital investments, and ecosystem strategies across the Middle East, North Africa, Southeast Asia, and Europe. His expertise covers AI digital infrastructure, hyperscale strategies, towers, non-terrestrial networks, as well as the full tech and digital investment life cycle—from strategy and playbook design to execution, value creation, and post-merger integration. He also plays an active role in corporate venture capital and venture building, helping launch digital ventures and investment platforms aligned with national and sectoral priorities.

Prateek Chauhan is a principal with Strategy& Middle East. Based in Dubai, he is a member of the technology, media, telecommunications, and digital practice in the Middle East. He has more than 11 years of consulting experience across India, the Middle East, and Southeast Asia. He specializes in AI and digital infrastructure, advising technology champions and digital infrastructure players on their strategic agenda spanning business transformation, growth, and commercialization.

Kirolous Zikry is a principal with Strategy& Middle East. Based in Riyadh, he is a member of the technology, media, telecommunications, and digital practice in the Middle East. He is a telecom and tech strategist with 17 years of experience leading digital infrastructure strategies and initiatives across global markets. He focuses on design, go-to-market, operations and growth of data centers players.

EXECUTIVE SUMMARY

The artificial intelligence (AI) revolution has reached a pivotal moment, and the Gulf Cooperation Council (GCC) countries[1] are key players in its next phase. The spotlight until recently has been on building ever-larger AI models. The focus is now shifting toward more efficient frontier models and decentralized architectures.

This fundamental reshaping of the landscape comes at a time when the GCC region has gained momentum as a hub for AI infrastructure. Recent landmark agreements on the AI and data center fronts during a U.S. presidential visit have highlighted the region's aim of transforming the Gulf into a global hub for AI and validated global investors' interest in the region. Countries like Saudi Arabia, the United Arab Emirates (UAE), and Qatar are using their energy wealth and strategic positioning to lay the groundwork for tomorrow's data centers and digital ecosystems. The convergence of technological shifts and regional ambition thus presents an opportunity for GCC countries to position themselves as a central force in global AI infrastructure.

For policymakers, the message is clear: GCC countries are doing more than closing the gap—they are building a new "digital Silk Road." To realize such ambitions, governments need to focus efforts to attract hyperscalers and investors into AI and digital infrastructure. They can do so by creating investor-friendly regulatory policy, offering incentives, putting in place ethical frameworks, facilitating smooth asset buildup, ensuring access to clean energy, strengthening regional connectivity, supporting hardware localization and talent development, and enabling a holistic AI ecosystem. For their part, private-sector players and regional data center developers should carefully adapt their positioning and offerings to tap into global and regional demand.

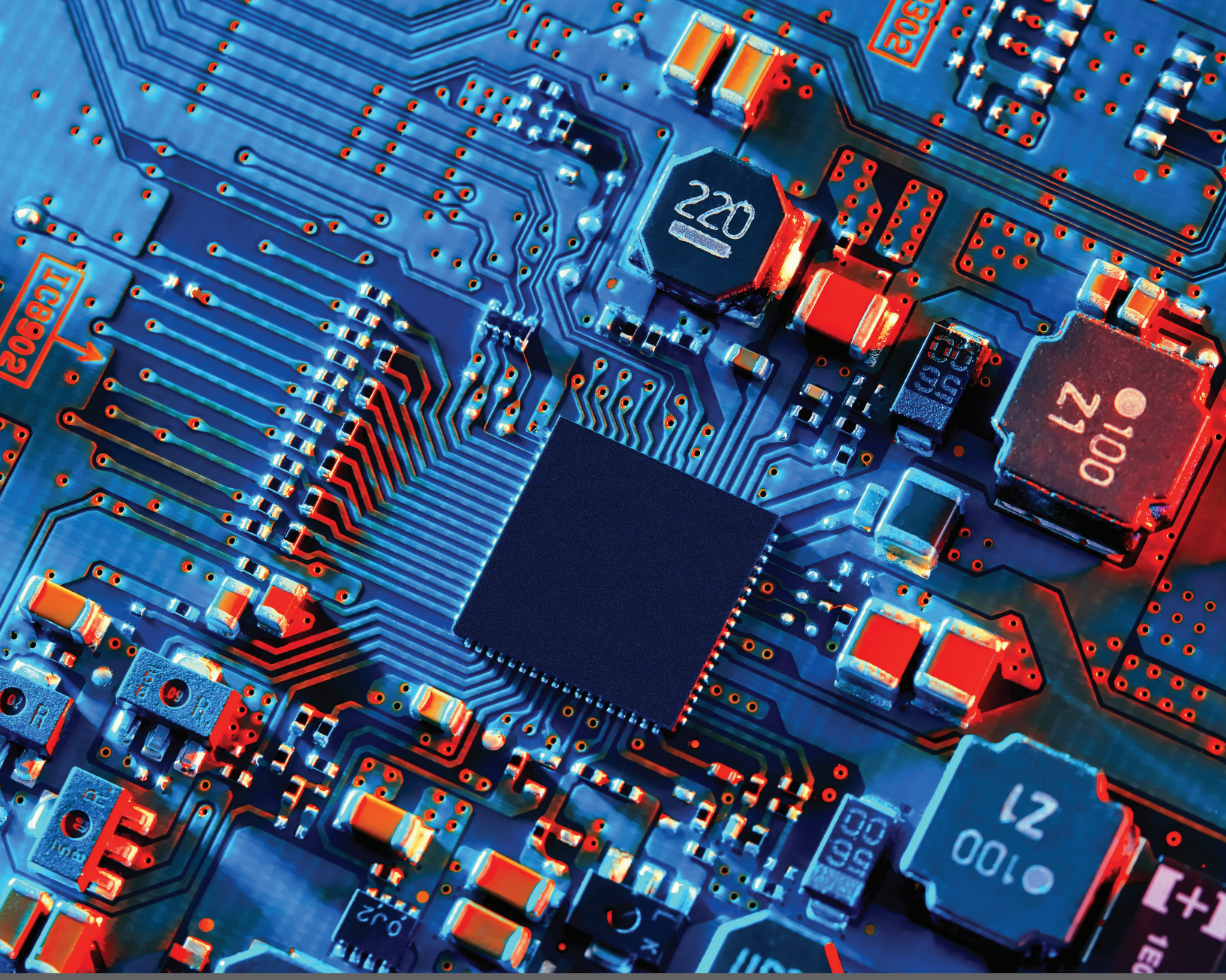
THE AI MODEL PARADIGM RESET PRESENTS AN OPPORTUNITY FOR GCC COUNTRIES

The reshaping of the AI landscape with the emergence of smaller, more efficient, and less expensive frontier models marks a paradigm reset. Among the main implications of this shift is that cutting-edge AI is no longer purely the domain of centralized global players with huge capital spending capabilities. Rather, it puts AI inferencing and training into the hands of new players in new geographies, including the GCC region if its countries can develop and train an ecosystem of fine-tuned models that would accelerate localized AI infrastructure development. The fact that AI is more distributed, more flexible, and more addressable from new geographies amounts to a major opportunity for GCC countries at a time when they are becoming significant global AI players.

The impact on cost of the paradigm reset is striking: DeepSeek's V3 costs as little as US$6 million for 14.8 trillion tokens, or only about 5 percent of the cost of OpenAI's GPT-4.[2] These models are thus much more cost-efficient for fine-tuning models and inferencing. The decline in inference costs, in particular, could accelerate growth in inferencing workloads being deployed across diverse use cases and latency needs.[3] Moreover, the region can access about 50 percent of the inferencing use cases that require latency of more than 250 milliseconds (ms), according to Strategy& analysis. That represents an opportunity for the region to attract global AI inferencing workloads.

No less significant is the ease of accessing the technology required. Because the frontier models have a lesser need for powerful computing, they have a reduced reliance on high-end chips, which are hard to obtain, given the huge demand; and which in the recent past were still subject to trade restrictions, and can incur high capital expenditures. For example, DeepSeek's V3 model needed only 2,000 of the highly specialized Nvidia-made H800 chips to train its model, compared with the 16,000 chips of H100 or more needed by its competitors.[4]

GCC countries can also seize the opportunity to attract global AI training workloads to the region. Given that AI training is latency agnostic and requires massive infrastructure as well as access to energy, GCC countries can position themselves to compete globally for such workloads.



> GCC countries can seize the opportunity to attract global AI training workloads to the region.

GCC COUNTRIES POSSESS COMPETITIVE ADVANTAGES

With its strategic geographic location, ambitious national AI strategies, access to capital, and growing focus on sustainability and data sovereignty,[5] the GCC region is rapidly positioning itself as a global hub for AI infrastructure, to serve global inferencing and training needs beyond the regional needs. Key competitive advantages include favorable energy economics, a strategic location and low-latency international connectivity, accelerated local demand, and regional innovation. The recent announcement of a 5 gigawatts (GW) capacity AI data facility in Abu Dhabi, which will be one of the largest globally, is a testament to the region's competitive edge.[6]

Favorable energy economics

The GCC region has ample, ultra-low-cost energy. For instance, Saudi Arabia offers an average electricity cost of 4.8 U.S. cents per kilowatt-hour (kWh) and Qatar boasts a cost of just 3.6 U.S. cents per kWh, compared with average costs in the U.S. of 18 U.S. cents per kWh and in the United Kingdom of 41 U.S. cents per kWh. GCC countries have the world's lowest solar energy costs and abundant land on which to build renewable generating capacity. Indeed, six of the ten lowest-cost existing solar projects globally are in the GCC region, including Saudi Arabia's Al Shuaibah project, which can generate solar energy at a levelized cost of electricity of 1.04 U.S. cents per kWh, as well as Abu Dhabi's large solar energy plant, which is currently under construction.[7] Such projects position the region as an attractive destination for sustainable data center growth, despite some power usage effectiveness challenges presented by the hot climate.

Strategic location and low-latency international connectivity

The GCC region also has a strong advantage in its location and international connectivity possibilities. With more than 30 subsea cable systems landing across Saudi Arabia, the UAE, Oman, and Qatar, the GCC offers a dense, low-latency, high-bandwidth crossroads between Europe, Asia, and Africa that perfectly suits globally oriented AI data center deployments.[8] Most enterprise AI workloads do not require real-time processing, and given its proximity to Europe (which has a latency of about 100 ms), and Singapore (about 80 ms), the GCC could serve up to 60 percent of global inferencing demand, acting as a gateway to the world, according to Strategy& analysis.

Accelerated local demand

The GCC region is spearheading several initiatives to catalyze AI adoption, data center localization, and cloud migration.

Ambitious government-led digital transformation initiatives and AI strategies and programs, such as Saudi Arabia's National AI Strategy and Qatar's National AI strategy, are driving AI adoption across private-sector and government entities.[9]

Implementation of stricter data localization and protection laws, such as the Personal Data Protection Law in Saudi Arabia, is driving the decentralization of data storage and computing infrastructure.[10] The push for sovereign AI,[11] focusing on developing national AI infrastructures, is creating strong demand for regional AI-centric data center infrastructure.

Traditional data center capacity is expected to remain relevant, as cloud adoption accelerates in GCC countries like Saudi Arabia and Qatar. According to a recent Strategy& survey, 68 percent of Middle East companies plan to migrate a majority of their operations to the cloud within the next two years.[12]

Availability of ample and relatively cheaper land

The region has ample land that can be made available for data centers. Moreover, land costs in the GCC are substantially lower than those in leading global data center hubs. For example, industrial land in Saudi Arabia typically ranges from $10 to $50 per square meter, compared to $150 to $600 per square meter in U.S. locations such as Northern Virginia. Although the land lease cost does not contribute significantly to the operating expenditure of a data center, this cost advantage further enhances the economic viability of constructing and scaling large-scale data center facilities.

Regional innovation and access to patient capital

Innovation in the region, from entrepreneurs and a startup ecosystem, is spurring AI adoption. The region has seen an explosion of AI-specific accelerators and funding that provide the impetus, support, technical know-how, and patient capital required to spur ideation in the space and drive the demand for national AI infrastructure. The UAE, for example, ranks ninth globally in AI enterprise density, according to the IFF Global Artificial Intelligence Competitiveness Index Report.[13] Saudi Arabia via its sovereign wealth fund is also launching a $40 billion AI-focused investment fund aiming to back AI startups, infrastructure, and GCCN–U.S. AI joint ventures, another testament of the ambition at hand.[14]

HOW GCC GOVERNMENTS CAN BUILD A ROBUST AI ECOSYSTEM

The GCC region has the potential to achieve 3 to 3.5 gigawatts in installed data center capacity by 2030, of which about one-third is expected to come from attracting global training and latency-agnostic inferencing workloads to the region. To achieve that potential, the region must undertake five steps that attract hyperscalers and investors to AI and digital infrastructure.

1

Putting in place investor-friendly regulatory policy, incentives, and data center special economic zones

To further accelerate data center growth, governments may want to introduce data center–specific zoning regulations, tax incentives, and land allocation preferences that encourage investment in the sector. Establishing a robust legal framework tailored to digital ecosystem players provides the regulatory clarity necessary to attract both domestic and international investors. A clear and consistent data governance and licensing regime is foundational to reducing perceived risk and building long-term investor confidence. In 2023, Saudi Arabia established the Cloud Computing Special Economic Zone, designed to attract foreign investment by offering a range of fiscal incentives and streamlined regulatory procedures.[15] Similar dedicated zones tailored specifically to the needs of AI data center investors and operators could play a pivotal role in accelerating the development of the sector.

2

Facilitating investments throughout the asset buildup life cycle

The region needs to not only attract investments but also facilitate them by enabling investors throughout their investment life cycle as they build up assets. The region needs to improve investors' experiences by making it easy for them to gain access to incentives, permits, land, utilities, and connectivity. Countries such as Malaysia and Singapore, and U.S. states such as Virginia, have set up concierge-like offices that act as a single point of contact for global investors.[16]

3

Enabling access to affordable and clean energy

Access to affordable and reliable clean energy is paramount. Initiatives such as permitting renewable energy generation for isolated "island mode" data centers[17] and developing regulatory frameworks for energy distribution are critical in realizing data center growth. The region should establish a mechanism that enables 24 hours a day matching of energy demand with carbon-free energy sources, using real-time tracking and granular energy certificates. It should also fast-track grid access for data centers and related facilities by introducing dedicated approval channels with clear, enforceable time lines. Upgrading transmission capacity in priority zones designated for digital infrastructure is essential to prevent bottlenecks and align grid development with projected demand.

4

Ensuring robust connectivity
Although the GCC region has a robust international connectivity infrastructure, collaborative efforts between governments and private-sector stakeholders are required to diversify subsea cable routes and create redundancies as well as improve regulatory frameworks to ensure carrier neutrality and equitable pricing for wholesale fiber capacity.

5

Building a holistic ecosystem from talent to hardware
Governments and industry stakeholders need to focus on building a holistic AI ecosystem. This includes localizing chips and hardware supply chains and promoting workforce development to accommodate the rapid construction and scaling of data centers, and providing access to specialized engineering, procurement, and construction and operations and maintenance capabilities. Enabling the development of AI platforms and applications, accelerating AI adoption, and maintaining the availability of funding for research and development are also a must.

> Providers should evaluate their strategic positioning—whether as hyperscaler enablers, national champions, or niche service providers—and adapt their offerings to meet global, regional, and addressable demand.

THE OPPORTUNITY FOR PRIVATE-SECTOR PLAYERS, INCLUDING HYPERSCALERS AND INVESTORS

Regional private players hoping to make significant investments in AI-focused data centers need to define a clear and differentiated value proposition to capitalize on growing demand.

Providers should evaluate their strategic positioning—whether as hyperscaler enablers, national champions, or niche service providers—and adapt their offerings to meet global, regional, and addressable demand. Key questions to consider include the types of workloads to offer: high-performance computing for AI training, AI inferencing, or hybrid models? Also, which customer segments drive demand for data centers in a given area: governments, hyperscalers, enterprises, or a combination?

Decisions about the scale and type of data center—central facilities, metro edge setups, or localized device edge infrastructures—also need to be made.[18] Should providers look to expand into AI-as-a-Service offerings such as GPUaaS?[19] Answering these questions allows data center providers to tailor their strategies and build a competitive edge in this rapidly evolving market. It is also important to use public–private partnerships. Partnerships between Saudi Arabia's Public Investment Fund and such companies as Google Cloud,[20] and G42 and Microsoft,[21] indicate the need for companies to work closely with national and government (or semi-government) entities to drive regional growth.

BUILDING A DIGITAL SILK ROAD

Ultimately, AI isn't just a technology play—it's a geopolitical one. AI is becoming a pillar of national competitiveness, and data infrastructure is its foundation. For GCC countries, investing in data centers can be the key to leapfrogging competitors. The original Silk Road connected the East and West through trade. Today's equivalent will run on fiber, graphics processing units, and AI computing. GCC countries sit at the heart of this route—not just as a transit point, but as a destination. If governments and private players align on vision, strategy, and execution, the region could define the future of AI infrastructure.

ENDNOTES

1. The GCC countries are Bahrain, Kuwait, Oman, Qatar, Saudi Arabia, and the United Arab Emirates.

2. For cost comparisons, see DeepSeek-V3 Technical Report (https://arxiv.org/pdf/2412.19437); and Will Knight, "OpenAI's CEO Says the Age of Giant AI Models Is Already Over," *Wired*, April 17, 2025 (https://www.wired.com/story/openai-ceo-sam-altman-the-age-of-giant-ai-models-is-already-over/). A token is a unit of data that an AI model processes.

3. Latency need refers to the delay involved in data transmission or processing.

4. Jess Weatherbed, "China's DeepSeek AI Is Hitting Nvidia Where It Hurts," The Verge, January 31, 2025 (https://www.theverge.com/2025/1/27/24352801/deepseek-ai-chatbot-chatgpt-ios-app-store); Justine Calma, "AI Is 'an Energy Hog' but DeepSeek Could Change That," The Verge, January 31, 2025 (https://www.theverge.com/climate-change/603622/deepseek-ai-environment-energy-climate); Cade Metz and Meaghan Tobin, "How Chinese A.I. Start-Up DeepSeek Is Competing with Silicon Valley Giants," *The New York Times*, January 23, 2025 (https://www.nytimes.com/2025/01/23/technology/deepseek-china-ai-chips.html).

5. Data sovereignty is the requirement to store data locally.

6. Data from Gulf Data Centre Association (https://gulfdca.com/en/5-things-to-know-about-the-gcc-data-centre-market/). U.S. Data Center Power Market Landscape Report 2024-2029, GlobeNewswire (https://tinyurl.com/yamn49ec).

7. Abimanyu Ghoshal, "World's Largest Solar Energy Plant Begins to Take Shape in Abu Dhabi," New Atlas, January 22, 2025 (https://newatlas.com/energy/abu-dhabi-largest-solar-energy-plant-construction/).

8. Martin Atkinson, "Can the Arabian Peninsula Become a Global Digital Crossroads?" Equinix, January 21, 2025 (https://blog.equinix.com/blog/2025/01/21/can-the-arabian-peninsula-become-a-global-digital-crossroads/).

9. Ministry of Transport and Communications and Hamad Bin Khalifa University, "National Artificial Intelligence Strategy for Qatar" (https://tinyurl.com/36a3dany); Saudi Digital and AI Authority, "National Strategy for Data & AI" (https://tinyurl.com/22b2ekry).

10. "Saudi Arabia's New Personal Data Protection Law in Force," DLA Piper, February 19, 2024 (https://tinyurl.com/43brjzxd).

11. Sovereign AI is an AI stack (including data, infrastructure, and a domestically developed frontier model) deployed within a country.

12. Achilles Drettas, Nikolaos Lioulis, Rajat Chowdhary, and Mina Ghabbour, “Cloud Computing in the Middle East: New Opportunities for Companies and Cloud Providers,” Strategy&, 2023 (https://www.strategyand.pwc.com/m1/en/strategic-foresight/sector-strategies/technology/cloud-computing.html).

13. “IFF Global AI Competitiveness Index Report Highlights UAE’s Achievements in AI,” February 24, 2025 (https://www.ifforum.org/index.php/show/7430).

14. Maureen Farrell and Rob Copeland, “Saudi Arabia Plans $40 Billion Push Into Artificial Intelligence,” *The New York Times*, March 19, 2024 (https://www.nytimes.com/2024/03/19/business/saudi-arabia-investment-artificial-intelligence.html).

15. Communications, Space & Technology Commission, “Cloud Computing Special Economic Zone” (https://www.cst.gov.sa/en/services/licensing/Pages/Cloud-Computing-Special-Economic-Zone.aspx).

16. Malaysia Digital Economy Corporation (https://mdec.my/); Digital Industry Singapore (DISG) (https://www.disg.gov.sg/); Virginia Economic Development Partnership (https://www.vedp.org/).

17. An “island mode” data center has its own power source so that it is not dependent on the national electricity grid.

18. Central facilities are locations that contain all the necessary computing and data storage infrastructure. Metro edge setups contain computing and data processing capacity near to sources of data generation to reduce processing times and decrease latency. Localized device edge infrastructures are located at the farthest tip of networks, allowing for even faster processing and less latency.

19. Graphics Processing Unit as a Service (GPUaaS) allows people to access powerful cloud computing when they need it.

20. “PIF and Google Cloud to Create Advanced AI Hub in Saudi Arabia,” tahawultech.com, October 30, 2024 (https://www.tahawultech.com/news/pif-and-google-cloud-to-create-advanced-ai-hub-in-saudi-arabia/).

21. “Microsoft Invests $1.5 Billion in Abu Dhabi’s G42 to Accelerate AI Development and Global Expansion,” *Microsoft News*, April 15, 2024 (https://tinyurl.com/mwjardkm).

strategy&

Part of the PwC network

Strategy&

Strategy& is a global strategy consulting business uniquely positioned to help deliver your best future: one that is built on differentiation from the inside out and tailored exactly to you. As part of PwC, every day we're building the winning systems that are at the heart of growth. We combine our powerful foresight with this tangible know-how, technology, and scale to help you create a better, more transformative strategy from day one.

As the only at-scale strategy business that's part of a global professional services network, we embed our strategy capabilities with frontline teams across PwC to show you where you need to go, the choices you'll need to make to get there, and how to get it right.

The result is an authentic strategy process powerful enough to capture possibility, while pragmatic enough to ensure effective delivery. It's the strategy that gets an organization through the changes of today and drives results that redefine tomorrow. It's the strategy that turns vision into reality. It's strategy, made real.

www.strategyand.pwc.com/me

Read the latest Ideation Center insights

ideationcenter.com

Connect with Strategy& Middle East

twitter.com/strategyandme

linkedin.com/company/strategyandme

strategyand.pwc.com/me

Connect with Strategy&

twitter.com/strategyand

linkedin.com/company/strategyand

youtube.com/user/strategyand

© 2025 PwC. All rights reserved. PwC refers to the PwC network and/or one or more of its member firms, each of which is a separate legal entity. Please see www.pwc.com/structure for further details. Mentions of Strategy& refer to the global team of practical strategists that is integrated within the PwC network of firms. For more about Strategy&, see www.strategyand.pwc.com. No reproduction is permitted in whole or part without written permission of PwC. Disclaimer: This content is for general purposes only, and should not be used as a substitute for consultation with professional advisors.