International Finance Corporation
WORLD BANK GROUP

EMERGING MARKET INSIGHTS
Published by DEC Private Markets

Risks *and* Opportunities *in* Data Center *and* AI Investment *in* Emerging Markets

BY DILEK AYKUT, KYUSAM JUNG, SELIN OZYURT, AND ESTEFANIA VERGARA COBOS

- Emerging markets and developing economies are becoming increasingly important locations for data-center investment as global cloud providers expand infrastructure to support artificial intelligence (AI) and digital services.
- China already hosts one of the world's largest data-center ecosystems, although much of its expansion has been driven by domestic technology firms and state-owned operators rather than foreign investment.
- By contrast, several other emerging markets and developing economies (EMDEs)—particularly in Asia and parts of Latin America—are attracting a growing number of projects from global hyperscalers and infrastructure investors.
- AI workloads are highly energy-intensive; global data center electricity demand is projected to nearly double by 2030. In EMDEs, without coordinated grid upgrades and renewable integration, AI growth could generate electricity shortages, higher energy costs, and undermine energy security and climate objectives.
- AI infrastructure investment can be a powerful development catalyst—but gains are not automatic. Productivity and employment effects depend on strong digital infrastructure and investment in other complementary sectors such as energy and transport, skilled labor, sound governance, and regulatory readiness that enable broad adoption across firms and institutions.
- Coordinated policies aligning digital expansion with energy planning and grid investment will be essential to maximize benefits while managing risks.

JULY 2026

Data Center and AI Investment Landscape in EMDEs

As generative AI expands rapidly, countries' ability to participate in the emerging AI-driven economy is increasingly shaped by national readiness and infrastructure capacity. For developing economies, making AI broadly accessible depends on several interrelated factors: physical and digital infrastructure, workforce skills, data availability and quality, financial resources, and supportive regulatory frameworks. Among these, infrastructure has emerged as the critical foundation. Access to computing power, data centers, reliable electricity, and high-speed connectivity increasingly determines whether countries and firms can participate meaningfully in the AI-driven economy.

Data centers are at the core of AI infrastructure, providing the computing power, storage, and high-speed networking required to train and run AI models. These processes typically depend on large clusters of specialized processors that operate in parallel and are connected through ultra-fast internal networks, rather than relying on a single machine (see *Appendix 1* for a technical overview).

While hyperscale cloud providers like Microsoft, Amazon, and Google drive the primary demand for AI computing capacity, the physical development of data centers involves a broader ecosystem of private actors. Specialized colocation providers, such as Equinix and Digital Realty, act as the primary developers, acquiring land, securing power permits, and building the essential power and cooling infrastructure. Hyperscalers then lease this capacity or enter into co-development agreements to deploy their hardware.

Beyond these operators, the data-center value chain spans multiple segments. Upstream includes semiconductor leaders (e.g., NVIDIA, TSMC) and server manufacturers (e.g., Foxconn, Dell). Infrastructure suppliers provide electrical and cooling solutions (e.g., Schneider Electric, Vertiv), while engineering firms and real-estate developers execute the construction, increasingly financed through specialized private credit markets and infrastructure funds.

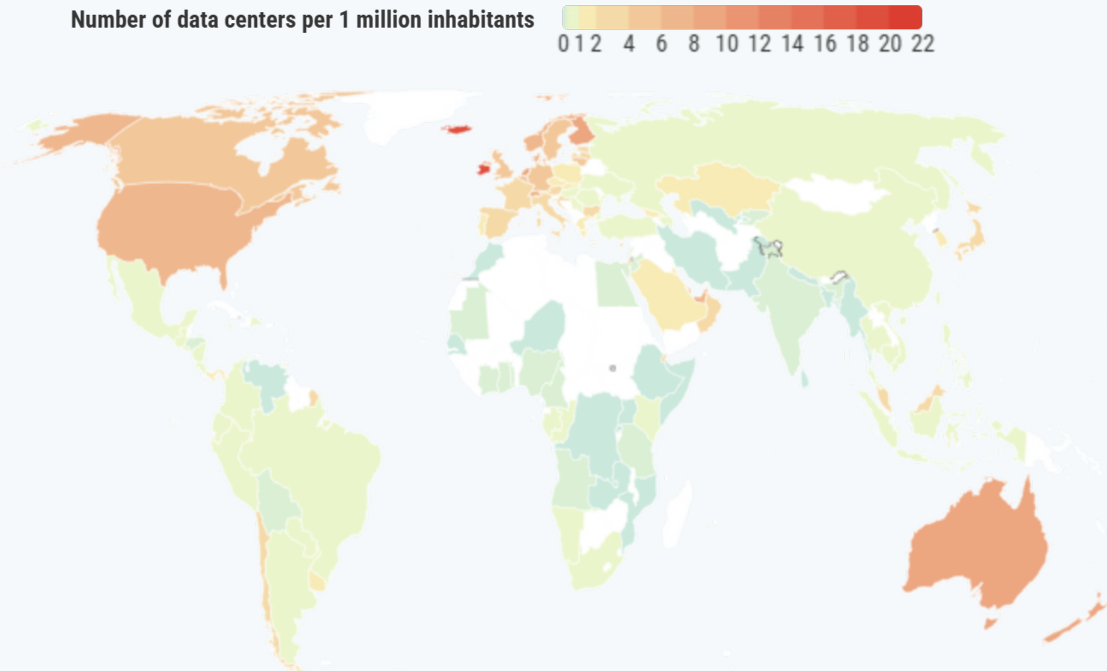
Data-center investments exhibit strong economies of scale. Total project costs rise with both facility size and installed power capacity, but larger developments distribute fixed costs—such as land acquisition, grid interconnection, and cooling infrastructure—more efficiently. Installed power capacity emerges as the main driver of capital expenditure shaped more by investment in computing intensity, power delivery systems, and cooling requirements rather than by building size alone.

The cost of building data centers also differs across income groups. In high-income countries, the cost of expanding data center capacity tends to increase more gradually than in upper-middle-income countries, possibly due to better access to technology and stronger economies of scale (*Appendix 2*). Adjusted for population, data centers are highly concentrated in North America and other advanced economies (*Figure 1*).

Nevertheless, EMDEs have become an increasingly important destination for data-center investment in recent years, as global cloud providers race to expand their footprint in fast-growing digital economies. Almost 3,000 data centers have been built across emerging markets and developing economies since the year 2000, with over half constructed in just the past few years. Several recent investment announcements by hyperscalers—

FIGURE 1

Geographical Distribution of Data Centers, 2025



Source: Straub et al. 2026. Infrastructure Foundations: From Current Assets to Future Growth. Sustainable Infrastructure Series. Washington, DC: World Bank.

including Google, Microsoft, Amazon Web Services, Nvidia, and OpenAI—across EMDEs highlight the sector's growing expansion and job creation potential (*Appendix 3*). The surge has been mirrored by a sharp rise in investment, which climbed from roughly $6bn in 2015 to more than $25bn by 2026, reaching a peak of about $31bn in 2024 (*Figure 2*).

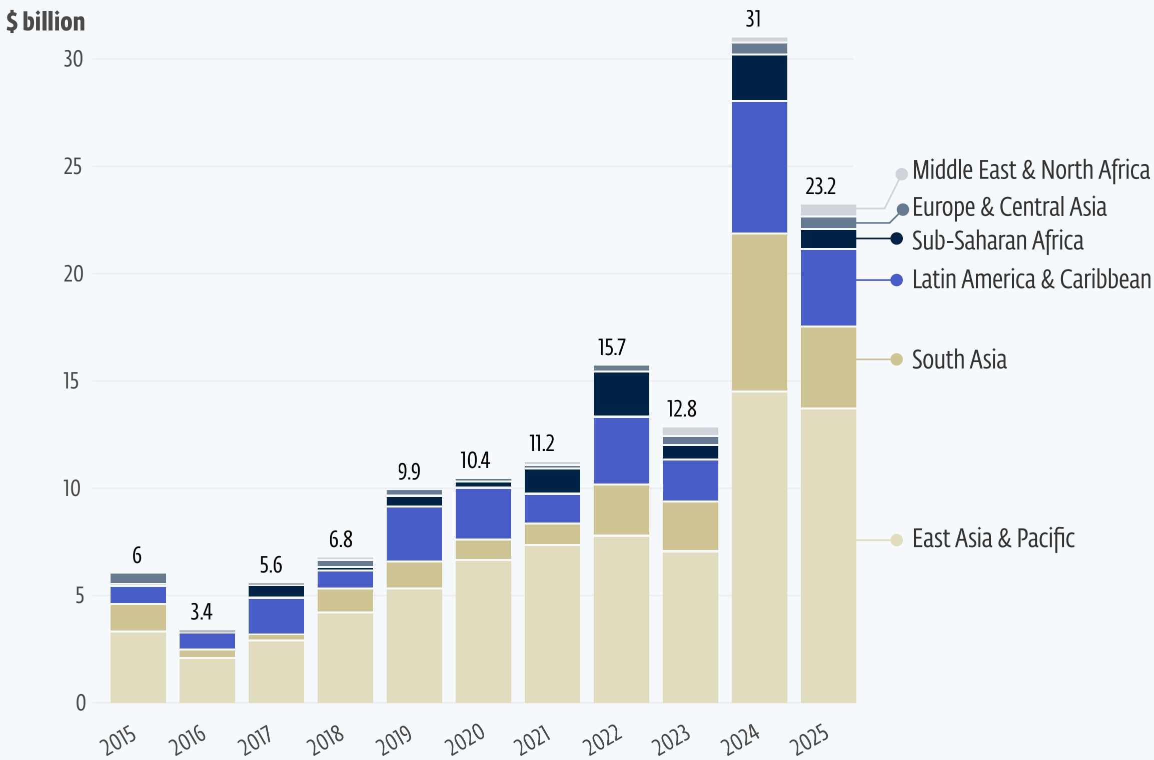
The majority of new data centers have been built in Asia. The region consistently accounts for the largest number of projects and has accelerated markedly since 2020, led by China, Malaysia and Indonesia. Latin America and the Caribbean and South Asia have also seen rapid increases, particularly after 2022, with Brazil and Mexico emerging as major regional hubs and Chile and Colombia developing as secondary locations. In the Middle East, the United Arab Emirates is positioning itself as a regional anchor for AI-related infrastructure.

By contrast, activity in Emerging Europe and Central Asia, the Middle East and North Africa, and Sub-Saharan Africa remains more limited in scale, with only sporadic project development overall. Nevertheless, countries such as South Africa, Kenya, Nigeria, and Morocco are attracting increasing investment despite persistent constraints related to power reliability and connectivity.

The distribution of data centers across EMDEs reflects a broader set of fundamentals. Data-center investment tends to concentrate in locations that combine reliable electricity supply, strong backbone fiber networks and international connectivity, available land and permitting pathways, access to finance, and a predictable regulatory

FIGURE 2

Data Center Investment in EMDEs, 2015–2026



Source: Straub et al. 2026. Infrastructure Foundations: From Current Assets to Future Growth. Sustainable Infrastructure Series. Washington, DC: World Bank.
Note: Values reflect reported project costs associated with data center construction and expansion and are used as a proxy for investment activity in non-high-income countries. Reported values combine observed project costs with estimated values where project level investment figures were not publicly available. Figures for recent years include announced or planned projects.

environment, particularly regarding data governance and cross-border data flows. Climatic conditions can also influence location decisions, as cooler average temperatures reduce cooling costs for large facilities.

Market fundamentals are equally important. Strong demand from enterprises and consumers, proximity to major economic hubs, and the presence of anchor customers such as hyperscale cloud and telecommunications operators all increase the attractiveness of potential locations. Policy choices can further shape investment patterns, increasingly through sovereign AI strategies and data-residency requirements that encourage local hosting of compute and data.[1,2]

The global surge in data-center investment is also reflected in FDI announcements. According to fDi Markets, announced investment in the sector increased sharply in recent years, reaching over $320 billion in 2025, up 74 percent

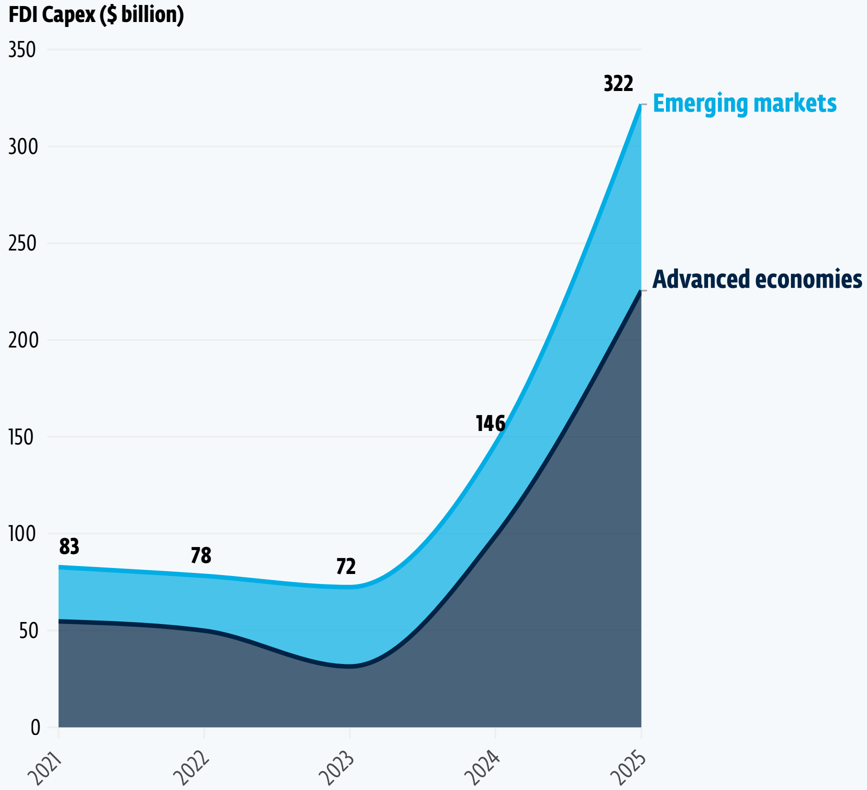
1 Malaysia's Digital Investment Office, which streamlines permitting and supports digital-economy investors, and Indonesia's OSS-RBA single-window system, which consolidates approvals and provides a clearer regulatory pathway for data-center investments, illustrate how targeted frameworks can attract large hyperscalers.

2 Sovereign AI refers to a national strategy to control the AI stack via domestic infrastructure and data. Key initiatives include India's $1.2bn GPUcompute mission, the UAE's Arabic-centric Jais and Falcon models, and Brazil's $4bn AI plan for public services and localized infrastructure.

year-on-year and accounting for nearly a quarter of global Foreign Direct Investment (FDI) capital expenditure (*Figure 3*). Although advanced economies capture most investment value, EMDEs account for roughly 40 percent of announced projects, highlighting their growing role in the global expansion of data-center infrastructure.

FIGURE 3

Global FDI Announcements into Data Centers



Source: fDI Markets and authors' calculation

Nevertheless, EMDE destinations of FDI announcements (*Appendix 3*) also reflect the concentration of investment within this group of countries. Malaysia, India, and Brazil stand out as the main recipients of data-center FDI. The scale of some of these commitments is illustrated by Microsoft's $17.5 billion data-center investment in India, announced in December 2025 and expected to be implemented by 2029. By contrast, China records relatively limited FDI announcements, as most data-center investment there is driven by domestic technology firms. In Africa, Nigeria and South Africa have each attracted more than $5 billion since 2020. As many projects were announced only recently and require several years to build, much of the associated investment and economic impact will materialize gradually in the coming years (and, in some cases, may be delayed or halted before fruition).

In terms of investor origin, most data-center FDI announcements are led by firms from North America, reflecting the dominant role of U.S.-based hyperscale cloud providers and technology companies in global digital infrastructure investment (*Table 1*). Firms from East Asia and Pacific and Western Europe also participate actively, often expanding within their own regions. Chinese firms have also become significant outward investors, accounting for about $54 billion in announced investments in 2025, with major projects in Brazil ($41 billion) as well as Thailand, Republic of Korea, and France. Interestingly, despite persistent infrastructure constraints, Africa attracted more data-center FDI projects than Emerging Europe, with African firms investing primarily within the region.

Beyond foreign capital inflows, domestic investment is also reshaping the geography of AI infrastructure across EMDEs. Many Asian EMDEs invest heavily in AI and cloud technologies, complementing foreign-led infrastructure expansion. China leads in

TABLE 1

Source and Destination of FDI Announcements into Data Center

2016 – September 2025 ($ billion)

DESTINATION \ SOURCE	North America	East Asia and Pacific	Western Europe	UAE	Africa	Other
Western Europe	199	25	64	45	-	3
East Asia and Pacific	133	92	14	3	-	2
North America	20	27	26	23	-	2
Latin America and Caribbean	35	6	2	-	-	6
Africa	13	4	5	2	8	1
Middle East	20	4	3	2	-	-
Emerging Europe	6	-	4	-	-	-
TOTAL	**426**	**158**	**118**	**75**	**8**	**13**

Source: fDi Markets

AI-related spending[3] as a share of GDP, followed by Malaysia, while cloud spending is particularly high in Thailand and Malaysia (*Appendix 4*). More broadly, Asian economies are developing distinct AI infrastructure strategies aligned with their roles in the global semiconductor value chain, gradually shifting from passive hardware suppliers toward specialized nodes within the global AI ecosystem.[4]

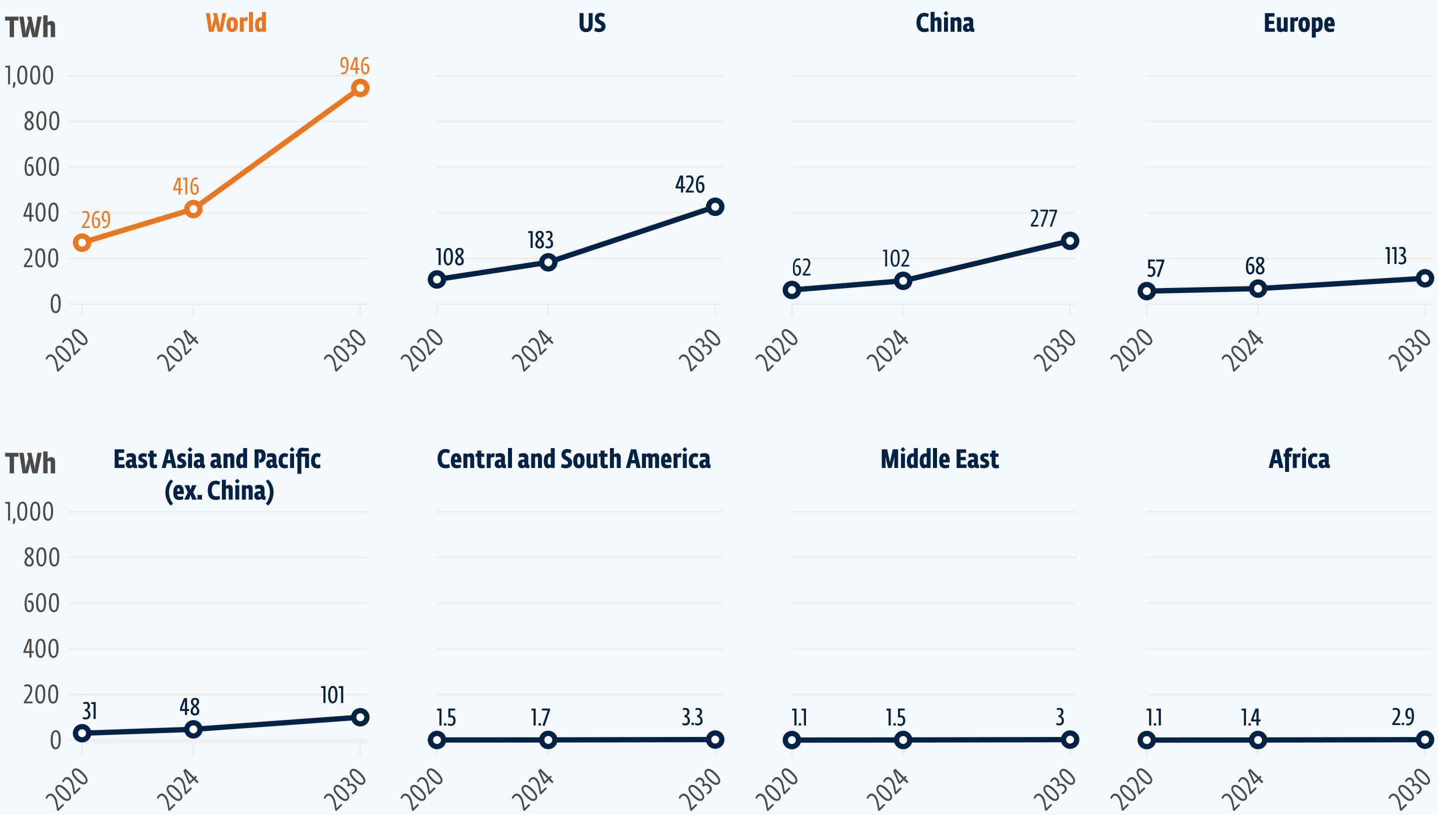
China illustrates the importance of domestic investment and policies in this sector. Despite hosting one of the world's largest data-center markets, it attracts little foreign FDI because most facilities are built and financed by domestic technology firms and state-owned telecom operators such as Alibaba Cloud, Tencent, Baidu, China Telecom, and China Mobile (Interesse 2023). Government initiatives like the Eastern Data Western Computing project have further accelerated expansion by shifting energy-intensive computing to western provinces with cheaper power, reinforcing China's role as a major data-center hub driven primarily by domestic investment (Chan et al. 2025).

3 AI-related spending includes hardware and software acquisition, fees paid to cloud service providers, and systems integration consultation exclude physical construction cost (IDC "Worldwide AI and generative AI Spending Guide").

4 Malaysia is leveraging its 13 percent global share in the Outsourced Semiconductor Assembly and Test (OSAT) market to transition into an "AI Factory" hub through advanced packaging and data centers based on Central Processing Unit (CPU). Thailand utilizes its cloud-first readiness and hardware base to develop "Sovereign AI" for localized services. Viet Nam is shifting toward high-value semiconductor design (Fabless), while the Philippines integrates AI infrastructure to boost productivity in its massive Business Process Outsourcing (BPO) sector, supported by its strong electronics export base.

FIGURE 4

Energy Consumption of Data Centers



Source: IEA (2025)

EMDE Challenges Arising from AI Investment[5] Expansion

AI deployment is rapidly reshaping global electricity demand, with data centers now among the fastest-growing power consumers. Operating continuously, requiring high-voltage supply, and relying on energy-intensive cooling for Graphics Processing Units (GPUs), these facilities have driven global data center electricity demand up by roughly 12 percent per year since 2017, more than four times faster than overall consumption growth (International Energy Agency 2025). The United States, China, and Europe account for about 85 percent of total demand. By 2030, global data center electricity use is projected to more than double to 946 TWh, led by the United States and China, with significant increases in other EMDEs (IEA 2025, *Figure 4*).

The climate and energy implications of this expansion, however, depend critically on the underlying generation mix, grid carbon intensity, and, where relevant, the marginal sources of

5 AI investment here is no longer limited to data centers; it encompasses the broader ecosystem of software, cloud service, and system integration. This shifts the focus from the 'physical container' to the economic value generated by AI utilization.

electricity during peak demand hours.

Countries with expanding AI activity face mounting pressure on their electricity systems, particularly where reserve margins are low, outages are frequent, and transmission networks remain weak. Meeting the continuous load requirements of hyperscale data centers requires substantial investment in substations, grid reinforcement, redundancy systems, and reliable baseload capacity, stretching utilities with limited operational space. Without complementary upgrades, rising AI-related demand could lead to more frequent outages, higher tariffs, and competition for reliable power between data centers and households or small enterprises.

Ensuring sustainable AI expansion will therefore require coordinated investments to strengthen grid capacity, modernize transmission infrastructure, and accelerate the integration of renewable energy sources. Some states in the United States have begun adjusting regulatory frameworks to manage these risks and require large electricity users such as data centers to contribute to grid infrastructure costs or secure dedicated power supply arrangements. Texas, for example, requires large consumers to participate in cost-sharing mechanisms for grid expansion, while Oregon has introduced a "large energy user" tariff class designed to prevent grid upgrade costs from being passed on to residential customers (S&P Global 2025; JLL 2026).

Empirical studies also suggest that rapid growth in AI-related electricity demand could influence energy prices if supply expansion does not keep pace. Estimates indicate that the combined impact of AI workloads, cryptocurrency mining, and broader electrification could raise U.S. wholesale electricity prices by roughly 19 percent between 2025 and 2028 (Chandramowli et al. 2024). Under scenarios where renewable expansion and grid investment lag behind demand growth, AI data-center expansion alone could increase retail electricity prices by roughly 8-9 percent by 2030 (IMF 2025). To prevent residential price spikes, grid upgrades can be paired with long-term electricity procurement contracts and peak-load incentives that encourage data centers to reduce consumption during high-demand periods (Burian et al. 2025).

AI-driven electricity demand in EMDEs risks reinforcing fossil-fuel dependence, particularly where energy systems are already constrained. Data centers are already disproportionately located in areas where the local renewable generation share is below the national average, implying closer reliance on thermal power in many sites. Meeting the additional power needs of data centers and high-performance computing hubs will depend heavily on each country's generation mix and available reserve capacity. In systems where renewable deployment and grid expansion lag behind digital demand, incremental load could be met in the short term by coal and gas.[6] In water-stressed economies, additional pressure on hydropower resources can further tighten supply, as lower reservoir levels reduce generation capacity and increase reliance on thermal alternatives.

In advanced economies, technology firms are increasingly exploring alternative energy sources to secure reliable low-carbon electricity for AI infrastructure such as nuclear power. For

6 In the US, AI server deployment is projected to emit 24–44 million metric tons of CO2_annually by 2030—equivalent to adding 5–10 million cars to the road (Xiao, T. et al., 2025). This underscores the disproportionate environmental transition risks facing EMDEs.

example, Microsoft has signed a long-term agreement supporting the restart of a nuclear power reactor at the Crane Clean Energy Center in the United States, while Amazon and Meta have invested in small modular reactor (SMR) projects intended to power future AI campuses. These developments highlight the scale and reliability requirements of next-generation AI infrastructure and may foreshadow similar energy–technology partnerships elsewhere.

Beyond electricity systems, AI infrastructure expansion also intensifies pressures along the water–energy nexus. Hyperscale data centers require substantial cooling, and in some cases can consume billions of liters of water annually. This poses an additional challenge in water-stressed EMDEs, particularly as next-generation high-density computing systems increasingly rely on liquid-cooling technologies. Although electricity remains the primary constraint, water availability and water-use efficiency are becoming increasingly important considerations as AI infrastructure scales.[7]

These dynamics risk shifting EMDEs into emissions pathways that are misaligned with their climate commitments and Nationally Determined Contributions (NDCs) targets.[8] In Southeast Asia, data-center expansion continues to rely predominantly on fossil fuel–based electricity systems, with operational emissions in Malaysia and Indonesia projected to increase severalfold by 2030 despite clean energy commitments (Ember 2025).

The environmental implications of AI infrastructure expansion are not limited to carbon emissions. Beyond increasing greenhouse gas emissions, the large-scale expansion of generation capacity, transmission networks, and cooling systems required to support AI infrastructure may also contribute to land-use change, biodiversity loss, and heightened water stress, particularly in river basins that are already under pressure (Zewe 2025).

Economic Opportunities and Spillovers from AI Investment in EMDEs

Data centers provide the physical foundation for AI, but its economic impact is driven by the software, services, and applications built on this infrastructure, which can catalyze productivity and growth across EMDEs.

AI-related infrastructure investment can be a driver of productivity and development in EMDEs by accelerating digital adoption, upgrading workforce skills, and fostering new economic activity. Katz et al. (2025b) find that greater investment in AI is associated with higher GDP, with estimated effects substantially larger than those linked to cloud investment alone. In 2024, AI spending and productivity spillovers across eleven Asia-Pacific economies are estimated to have generated approximately $247 billion, or 1.6 percent of their combined GDP, with most gains driven by productivity spillovers component.[9] Evidence from specific sectors further illustrates these gains. In agriculture,

7 Liquid cooling does not always increase total water consumption. In some designs, it can actually improve efficiency and reduce evaporation.

8 Nationally Determined Contributions (NDCs) are the climate-action commitments that countries submit under the Paris Agreement, outlining their targets for reducing greenhouse-gas emissions and adapting to climate impacts. Each country is expected to strengthen its NDC over time in line with global decarbonization goals.

9 Australia, India, Indonesia, Japan, Korea, Malaysia, New Zealand, Philippines, Singapore, Thailand, and Viet Nam.

logistics, and manufacturing, AI applications have improved forecasting, reduced waste, and streamlined supply chains, while AI-enabled irrigation and fertilization systems have raised yields by an estimated 5–15 percent (Hoque and Padhiary 2024; Nautiyal et al. 2025).

At the firm level, advanced digital technologies, including generative AI–enabled tools, can enhance firm-level productivity by improving access to data, analytics, automation, and digital services (OECD 2025). However, aggregate gains may be modest and uneven across sectors, depending on the range of tasks affected and the extent of cost savings achieved (Acemoglu 2025).

Labor market effects are mixed. Evidence suggests that each 1 percent increase in AI spending raises demand for skilled labor by 0.21 percent but reduces demand for unskilled labor by 0.15 percent (Katz et al. 2025a), intensifying risks of wage polarization where AI benefits skilled workers more while displacing others (World Bank 2025). Systematic reviews highlight that generative AI can both create new roles and displace existing ones, underscoring the need for adaptive policy frameworks (Salari et al. 2025).

Beyond direct productivity effects, AI investments can catalyze the broader digital ecosystem. Data centers can attract complementary activities such as software development, cybersecurity, data services, cloud operations, and AI research, while generating employment along the value chain—from construction and engineering during the build phase to increasingly skilled roles as ecosystems mature. By localizing inference capacity, these investments can lower latency for domestic fintech and logistics platforms, making real-time digital services more viable for local firms and potentially boosting firm-level productivity (AWS 2024).

Data-center development can also generate localized economic spillovers. During construction phases, large projects require significant engineering, design, and construction services, often employing thousands of workers at peak activity (McKinsey 2025). The supply chain extends to equipment manufacturers producing servers, transformers, backup generators, and cooling systems (IoT Analytics 2025). Local governments may also benefit from substantial fiscal revenues; for example, Loudoun County in Northern Virginia expects nearly $900 million in annual tax revenues from data-center operations in fiscal year 2025 (Lincoln Institute of Land Policy 2025). At the same time, these benefits can be uneven, as most employment effects occur during construction rather than ongoing operation, and land demand associated with data-center development may increase local real-estate prices (World Bank 2024) or generate community concerns related to noise and environmental impacts (Lincoln Institute of Land Policy 2025).

When powered by renewables, hyperscale data centers can act as anchor clients for utility scale solar and wind projects through long-term power purchase agreements, a mechanism identified by International Energy Agency (IEA) and International Renewable Energy Agency (IRENA) as critical for accelerating clean energy deployment in EMDEs. Infrastructure investments associated with these facilities, including upgrades to fiber backbones, substations, and transmission networks can also strengthen system reliability and connectivity for households and firms, and may also potentially create jobs. Local hosting of sensitive data also supports digital sovereignty objectives and reduces reliance on foreign servers—an increasing priority across Asia, Africa, and Latin America (Internet Society 2022, De Frietas 2025).

Realizing productivity benefits from AI capacity is not automatic. Infrastructure enables access, but impact depends on adoption across firms, households, and public institutions, which in turn requires affordability, digital skills, complementary investments, and a predictable regulatory and data governance environment.[10]

Conclusion

Policy readiness appears to play an important role in shaping how much of this potential is realized. Economies with advanced AI and cloud strategies—such as Singapore, Australia, Japan, and Korea—tend to record a larger overall economic impact of AI and cloud investment, including direct spending and indirect spillover effects across the economy, estimated at around 3–4 percent of GDP compared with less than 1 percent in countries such as Indonesia and Viet Nam (Katz et al. 2025b). Model simulations further suggest strengthening AI and cloud policy frameworks could raise technology spending by up to 2 percentage points of GDP in some economies, unlocking additional growth.

Taken together, AI infrastructure can become a powerful development catalyst—but only under the right conditions. Countries that align digital ambitions with energy-sector planning and grid modernization are emerging as preferred destinations for hyperscalers. Malaysia, through its National Energy Transition Roadmap, Brazil with an electricity mix that is more than 80 percent renewable, and the UAE, which is developing nuclear- and solar-powered AI hubs, illustrate how coordinated power-sector strategies can attract large-scale data center investment. By contrast, economies where transmission upgrades and energy reforms lag risk remaining peripheral to the next wave of AI expansion. In South Africa, chronic load shedding undermines reliability. In Viet Nam, transmission bottlenecks and the 2023 power shortages highlight how grid constraints can deter the massive and stable electricity demand required for next-generation AI infrastructure.

As AI applications increasingly extend into real time and operational settings, some use cases require lower latency connectivity and closer integration between digital systems and physical infrastructure. Without adequate digital and supporting infrastructure, existing gaps in adoption could limit the ability of some economies to capture emerging productivity gains and widen the gap between digitally advanced economies and those left behind. However, as AI shifts from heavy training to real-time inference, 'Small AI' and Edge Computing offer EMDEs a strategic leapfrogging opportunity. By running lightweight models on local devices, these economies can bypass grid bottlenecks and deliver AI value via existing mobile networks. ■

10 ADB (2025), OECD (2024) and World Bank (2024) find that AI and cloud adoption generate economic impact through productivity and efficiency gains, though benefits vary widely with countries' digital infrastructure, skills, data governance and regulatory readiness. GSMA (2024) highlights early productivity and service-delivery gains from AI use case in Africa, while noting that outcomes depend heavily on connectivity, data infrastructure and skills.

References

- Acemoglu, D., (2025) 'The Simple Macroeconomics of AI,' Economic Policy, vol 40(121), pages 13-58 and Calvino, F., Reijerink, J., & Samek, L., (2025), 'The Effects of Generative AI on Productivity, Innovation and Entrepreneurship,' OECD Artificial Intelligence Papers, No. 39, available at https://doi.org/10.1787/b21df222-en

- ADB 2025 "Leveraging Artificial Intelligence and Cloud Computing to Accelerate Growth in Asia and the Pacific" https://www.adb.org/publications/artificial-intelligence-cloud-computing

- Amazon Web Services (AWS). 2024. "Reduce conversational AI response time through inference at the edge with AWS Local Zones". Available at https://aws.amazon.com/blogs/machine-learning/reduce-conversational-ai-response-time-through-inference-at-the-edge-with-aws-local-zones/

- Burian, V., & Stalla-Bourdillon, A., (2025) 'The Increasing Energy Demand of Artificial Intelligence and Its Impact on Commodity Prices,' European Central Bank, Economic Bulletin, Issue 2/2025, available at https://www.ecb.europa.eu/press/economic-bulletin/focus/2025/html/ecb.ebbox202502_03%7E8eba688e29.en.html

- Chan, K., Smith, G., Goodrich, J., DiPippo, G., & Pilz, K. F., (2025) 'Full Stack: China's Evolving Industrial Policy for AI,' RAND Corporation, available at https://www.rand.org/pubs/perspectives/PEA4012-1.html

- Chandramowli, Shankar, Patty Cook, Justin Mackovyak, Himali Parmar, and Maria Scheller (2024) Power Surge: Navigating US Electricity Demand Growth. New York: ICF, available at https://www.ourenergypolicy.org/resources/power-surge-navigating-us-electricity-demand-growth/

- De Freitas, M. V., (2025) 'Digital Sovereignty and Data Colonialism: Shaping a Just Digital Order for the Global South,' Policy Center for the New South, Policy Paper No. 38/25, available at https://www.policycenter.ma/sites/default/files/2025-10/PP_38-25%20%28Marcus%20Vini%CC%81cius%20De%20Freitas%29.pdf

- EMBER (2025)"From AI to emissions: Aligning ASEAN's digital growth with energy transition goals" https://ember-energy.org/latest-insights/global-electricity-review-2025/

- Gelvanovska-Garcia, N., Maciule, V., & Rossotto, C. M., (2024) 'Advancing Cloud and Data Infrastructure Markets,' World Bank, available at https://documents.worldbank.org/en/publication/documents-reports/documentdetail/099052824071033398

- Gorey, J., (2025) 'Data Drain: The Land and Water Impacts of the AI Boom,' Lincoln Institute of Land Policy, Land Lines Magazine, available at https://www.lincolninst.edu/publications/land-lines-magazine/articles/land-water-impacts-data-centers/

- GSMA (2024) "AI for Africa: Use cases delivering impact" https://www.gsma.com/solutions-and-impact/connectivity-for-good/mobile-for-development/gsma_resources/ai-for-africa-use-cases-delivering-impact/

- Hoque, A., & Padhiary, M., (2024) 'Automation and AI in Precision Agriculture: Innovations for Enhanced Crop Management and Sustainability,' Asian Journal of Research in Computer Science, vol. 17(10), pp. 95-109, available at https://journalajrcos.com/index.php/AJRCOS/article/view/512

- Humeau, E., & Deshpande, T., (2024) 'AI for Africa: Use Cases Delivering Impact,' GSMA, available at https://www.gsma.com/solutions-and-impact/connectivity-for-good/mobile-for-development/wp-content/uploads/2024/07/AI_for_Africa.pdf

- IDC, (2024, 2025) 'Worldwide AI and Generative AI Spending Guide,' International Data Corporation

- Interesse, G., (2023) 'China's Data Center Sector: Industry and Regulatory Insights,' China Briefing, available at https://www.china-briefing.com/news/chinas-data-center-sector-industry-and-regulatory-insights/

- International Energy Agency, (2025) 'Energy and AI,' IEA, Paris, available at https://www.iea.org/reports/energy-and-ai
- International Monetary Fund, (2025) 'Power Hungry: How AI Will Drive Energy Demand,' IMF Working Paper, available at https://www.imf.org/-/media/files/publications/wp/2025/english/wpiea2025081-print-pdf.pdf
- International Renewable Energy Agency, (2018) 'Corporate Sourcing of Renewables: Market and Industry Trends – REmade Index 2018,' IRENA, Abu Dhabi, available at https://www.irena.org/-/media/Files/IRENA/Agency/Publication/2018/May/IRENA_Corporate_sourcing_2018.pdf
- Internet Society, (2022) 'Navigating Digital Sovereignty and Its Impact on the Internet,' Internet Society, available at https://www.internetsociety.org/wp-content/uploads/2022/11/Digital-Sovereignty.pdf
- IoT Analytics (2025) "Data center infrastructure market: AI-driven CapEx pushing IT and facility equipment" https://iot-analytics.com/data-center-infrastructure-market/
- JLL, (2026) '2026 Global Data Center Market Outlook: Navigating AI Demand, Power Constraints and Global Opportunities in 2026,' JLL Research, available at https://www.jll.com/en-us/insights/market-outlook/data-center-outlook
- Katz, R., Jung, J., Beschorner, N., Park, D., Beirne, J., & Kong, B., (2025a) 'Leveraging Artificial Intelligence and Cloud Computing to Accelerate Growth in Asia and the Pacific,' Asian Development Bank, ADB Briefs No. 352, doi: 10.22617/BRF250337, available at https://www.adb.org/publications/artificial-intelligence-cloud-computing
- Katz, R., Jung, J., & Callorda, F., (2025b) 'Economic Impact of Artificial Intelligence and Cloud Computing Development in Asia and the Pacific,' Telecom Advisory Services, available at https://www.teleadvs.com/economic-impact-ofartificial-intelligence-and-cloud-computingdevelopment-in-asia-and-the-pacific/
- Lincoln Institute of Land Policy 2025 "Data Drain: The Land and Water Impacts of the AI Boom" https://www.lincolninst.edu/pt-br/publications/land-lines-magazine/articles/land-water-impacts-data-centers/
- Lueth, K. L., (2025) 'Data Center Infrastructure Market: AI-Driven CapEx Pushing IT and Facility Equipment Spending Toward $1 Trillion by 2030,' IoT Analytics, available at https://iot-analytics.com/data-center-infrastructure-market/
- McKinsey (2025) "Scaling bigger, faster, cheaper data centers with smarter designs" https://www.mckinsey.com/industries/private-capital/our-insights/scaling-bigger-faster-cheaper-data-centers-with-smarter-designs
- Nadhila, S., (2025) 'From AI to Emissions: Aligning ASEAN's Digital Growth With Energy Transition Goals,' EMBER, available at https://ember-energy.org/app/uploads/2025/05/Report-From-AI-toemissions-ASEAN.pdf
- Nicoletti, G., (2024) 'The Impact of Artificial Intelligence on Productivity, Distribution and Growth: Key Mechanisms, Initial Evidence and Policy Challenges,' OECD Artificial Intelligence Papers, doi: 10.1787/8d900037-en, available at https://www.oecd.org/en/publications/the-impact-of-artificial-intelligence-on-productivity-distribution-and-growth_8d900037-en.html
- OECD 2025 "The Effects of Generative AI on Productivity, Innovation and Entrepreneurship" https://www.oecd.org/content/dam/oecd/en/publications/reports/2025/06/the-effects-of-generative-ai-on-productivity-innovation-and-entrepreneurship_da1d085d/b21df222-en.pdfS&P Global Annual Report (2025) "Advancing Essential Intelligence" https://www.spglobal.com/en/annual-reports/2025
- Sachdeva, P., Pikul, P., Rauer, J., & Kok, E., (2025) 'Scaling Bigger, Faster, Cheaper Data Centers With Smarter Designs,' McKinsey & Company, available at https://www.mckinsey.com/industries/private-capital/our-insights/scaling-bigger-faster-cheaper-data-centers-with-smarter-designs

- Salari, N., Beiromvand, M., Hosseinian-Far, A., Habibi, J., Babajani, F., & Mohammadi, M., (2025) 'Impacts of Generative Artificial Intelligence on the Future of Labor Market: A Systematic Review,' ScienceDirect, available at https://www.sciencedirect.com/science/article/pii/S2451958825000673
- Seifi, N., & Ong, N., (2025) 'Reduce Conversational AI Response Time Through Inference at the Edge with AWS Local Zones,' AWS Blog, available at https://aws.amazon.com/blogs/machine-learning/reduce-conversational-ai-response-time-through-inference-at-the-edge-with-aws-local-zones/
- Straub, Stéphane, He He, Yue Li, Xinxin Lyu, Jevgenijs Steinbuks, Estefanía Vergara Cobos, Christopher Dann, Manuel García-Santana, and Harris Selod, (2026) 'Infrastructure Foundations: From Current Assets to Future Growth,' Sustainable Infrastructure Series, World Bank, Washington, DC, available at https://openknowledge.worldbank.org/entities/publication/c35564b6-20e5-4606-a9df-d5735f8481d3
- Thompson, D., Lenoir, T., & Liu, P., (2025) 'Geopolitics of Data Centers: An AI Showdown That Will Reshape the World,' S&P Global, available at https://www.spglobal.com/en/research-insights/special-reports/look-forward/data-center-frontiers/geopolitics-data-sovereignty-data-center-security
- World Bank (2024) "Advancing Cloud and Data Infrastructure Markets https://documents.worldbank.org/en/publication/documents-reports/documentdetail/099052824071033398
- World Bank, (2025) 'Digital Progress and Trends Report 2025: Strengthening AI Foundations,' World Bank, Washington, DC, doi: 10.1596/978-1-4648-2264-3, available at https://openknowledge.worldbank.org/server/api/core/bitstreams/f2509a0f-7153-4f32-b180-bc11e90c4940/content
- Xiao, T., Fuso Nerini, F., Matthews, H. D., Tavoni, M., & You, F., (2025) 'Environmental Impact and Net-Zero Pathways for Sustainable Artificial Intelligence Servers in the USA,' Nature Sustainability, vol 8, pages 1541-1553, available at https://www.nature.com/articles/s41893-025-01681-y
- Zewe, A., (2025) 'Explained: Generative AI's Environmental Impact,' MIT News, available at https://news.mit.edu/2025/explained-generative-ai-environmental-impact-0117

APPENDIX 1:
The Architecture of AI Systems:
Resource Requirements and Infrastructure Layers

AI systems operate through two distinct but complementary phases: training and inference. Model training is the resource-intensive phase in which models are developed and calibrated using large datasets, accelerators such as GPUs, high- bandwidth networking, storage, and substantial energy supply. Inference is the deployment phase where trained models generate outputs (such as text and images) repeatedly in response to user queries or automated processes. A single inference is cheaper than training, but billions of daily queries can make aggregate inference energy demand very large and even dominate total operation energy use over time. Understanding these requirements helps clarify why AI infrastructure is commonly organized around three interconnected layers:

1. **Compute and interconnect layer**: This layer consists of clusters of specialized processors such as GPUs connected by very fast communication links (such as NVLink), allowing many machines to work together to run AI models efficiently. For the cluster to operate as one system it requires high performance networking such as high-speed Ethernet. Due to this, large-scale model training is usually geographically centralized to capture economies of scale, secure power, and network capacity.
2. **Physical facilities**: Data centers[11] provide the space, electricity, cooling, and security needed to run AI workloads continuously. But not every facility can host high-density GPU clusters. AI capable sites, most often hyperscale or large regional facilities, are built for much higher electricity use per rack and more advanced cooling than traditional CPU-based data centers, because GPU workloads draw more power and generate more heat.[12]
3. **Transport and connectivity**: The network backbone, including submarine cables and last-mile connections, carries requests to where models run and returns outputs to users. The speed and reliability of these networks determine how quickly AI services can be accessed across regions and countries.

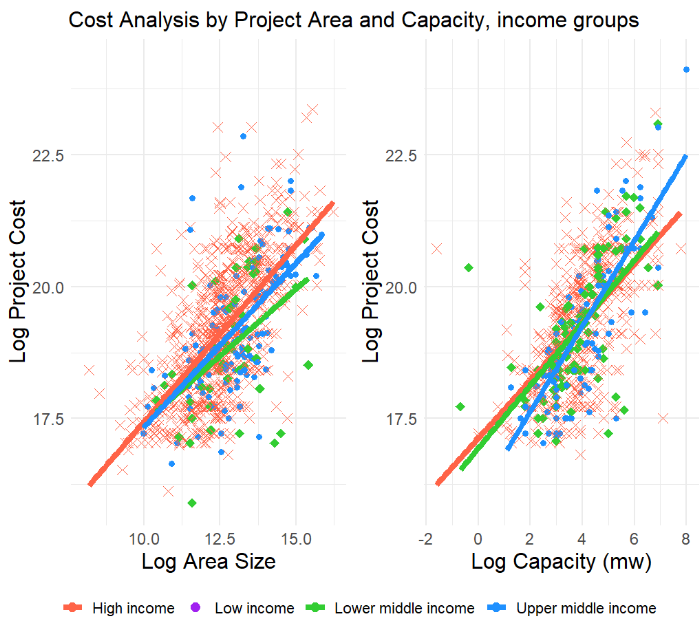
11 Data center demand is driven by both traditional IT and AI-driven workloads. According to McKinsey (2025), AI is the dominant driver, projected to account for $5.2 trillion (approx. 78 percent) of the $6.7 trillion global CAPEX and 70 percent of incremental capacity by 2030. This dominance is due to the higher capital intensity of AI infrastructure—requiring high-density GPU clusters and specialized cooling—compared to the CPU-based systems used for baseline digital services.

12 AI workloads refer to the computational tasks involved in training and deploying AI models. In this paper, the term covers three categories: (i) individual user queries (e.g., prompts to a language model), (ii) firm-level commercial applications (such as customer service automation, coding assistants, or document drafting), and (iii) backend AI processes embedded in digital services (including search, recommendation systems, fraud detection, or data analytics). These categories differ in scale, frequency, infrastructure intensity, and economic impact, and therefore have distinct implications for compute demand, network requirements, and development outcomes.

Layer	Stage	Function	Key Physical Infrastructure	Energy Intensity	Regional Pattern & EMDE Implication
Transport (Connectivity) "How data physically moves from users to data centers"	1. User Access	Connects end-users/ devices to the internet (Last-mile)	5G Towers, Fiber to-the-Home (FTTH), Wi-Fi Equipment	High (low per device, but large in total)	Critical Bottleneck. Weak last-mile connectivity in many EMDEs limits broad AI adoption.
	2. Long-distance Transport	Transmits data across borders and oceans (The Backbone).	Subsea Cables, Terrestrial Fiber	Low (Efficient optical transmission)	Strategic Asset. Cable landing countries (e.g., Brazil, Kenya, Indonesia) act as regional gateways, gaining leverage in digital divide.
	3. Traffic Exchange	Local hubs where networks connect to exchange data.	Internet Exchange Points (IXPs)	Low	Cost & Latency. Local IXPs are vital to keep data traffic domestic and reduce the cost of AI access for EMDE startups.
Compute (The "AI Factory")	4. Local Processing (Edge)	Fast processing near the user for real time process	Edge Data Centers, Content Delivery Network (CDN) Nodes (e.g., Cloudflare, Akamai; AWS Local Zones)	Medium (Cooling efficiency varies)	High Growth. Essential for real-time services (e.g., Smart Cities, instant language translation). Key investment target in Tier-2 cities.
"Where AI models are built and run"	5-A. Model Training (Development)	Creates the AI model. Requires massive parallel processing.	Hyperscale Data Centers, GPU Clusters with High Bandwidth Memory 4 (HBM4) (e.g., Nvidia (H100), AMD; Foundry: TSMC), Interconnect	Very High (Continuous maxload)	Centralized. Requires massive, continuous power and advanced grids (e.g., US, Nordics). Limited opportunity for most EMDEs.
	5-B. Model Inference (Execution)	Generates answers for users using pre-trained models.	Regional Data Centers, Accelerated Servers (e.g., Supermicro, Dell; Clouds: CoreWeave)	High (Variable load, 24/7 availability)	Regional/Distributed. Must be close to users. Primary driver of data center demand in Asia & LATAM.
Data & Software (Storage & Management) "Where data is stored and systems are managed"	6. Data Storage	Stores data and feeds it to GPUs at ultra-high speed.w	High-Speed Storage Systems (SSD) (e.g., Samsung, SK Hynix)	Medium (storage is less energy-intensive than computing, but runs continuously)	Sovereignty Driver. Data localization laws force storage infra to be built within national borders.
	7. Management Software	Manages workloads and GPU resources (Software layer).	GPU Management Platforms	Negligible (Software-defined)	Skill Gap. Requires highly skilled DevOps engineers, often scarce in EMDE markets.

Source: various sources, including IEA (2025), Nvidia, and IBM

APPENDIX 2:
Cost of Data Centers



Source: Straub et al. 2026. Infrastructure Foundations: From Current Assets to Future Growth. Sustainable Infrastructure Series. Washington, DC: World Bank.

APPENDIX 3:
Recent Data infrastructure Investment Announcement (2024~)

	Country	Company	Investment Size	Project Description	Impact (Jobs/GDP)
ASIA	**Malaysia**	Google	$2bn (2024)	First Google Data Center & Google Cloud Region	~26,500 jobs and $3.2 billion GDP contribution by 2030
		Microsoft	$2.2bn (2024)	Cloud & AI infra investment; AI Center of Excellence; national AI skills program	Training for ~200,000 people
		AWS	$6.2bn (to 2038)	Launch of AWS Asia Pacific Region with three Availability Zones	~3,500 jobs per year
		NVIDIA + YTL Power	$4.3bn (plan announced around 2023)	GPU-based AI Factory / Supercomputing Data Center using renewable energy	
	Indonesia	Microsoft	$1.7bn (2024)	Cloud and AI infra with AI skilling initiative	AI training for ~840,000 workers
	Thailand	Google	$1bn (2024)	Data Center and Google Cloud Region	~14,000 jobs per year (2025–29)
	India	Google	$15bn (2025–2030)	1 GW-scale AI data center & AI hub in Visakhapatnam (Andhra Pradesh)	~188,000 jobs
		AWS	$12.7bn (to 2030)	Expansion of Indian cloud regions incl. Hyderabad Region	~131,700 direct & indirect jobs
	Singapore	AWS	$9bn (to 2028)	Expansion of Asia-Pacific HQ and cloud capacity; new data-center clusters	~12,000 jobs and US$17.6 billion GDP impact
		Singtel + NVIDIA Singtel + NVIDIA	n/a (2024–25)	Regional AI Factory / GPU-as-a-Service hub for Singapore, Thailand, Indonesia.	

	Country	Company	Investment Size	Project Description	Impact (Jobs/GDP)
AFRICA	**Kenya**	Microsoft + G42	$1bn (announced 2024)	1 GW green data center campus to host East Africa's first Azure cloud region	supports last-mile internet for 20m people in Kenya and 40m across East Africa, plus AI/digital skills training
	South Africa	Microsoft	~ $300mn (2025-27)	Expansion of hyperscale cloud and AI infrastructure, including new data-center capacity and sustainability initiatives, and a digital-skills program	train 1 million South Africans by 2026, improving employability and strengthening the local AI ecosystem
		Visa	$57mn (2025–28)	Visa's first data center in Africa to localize payment processing, improve latency system resilience	
LATIN AMERICA	**Uruguay**	Google	$850mn (2024–)	Google's second Latin American data center, improving cloud reliability, regional AI compute availability, and subsea-cable capacity	
	Mexico	CloudHQ	$4.8bn (2025)	One of LATAM's largest planned DC campuses (six facilities, ~900 MW), designed to meet accelerating AI and cloud demand and support nearshoring trends	generate 7,200 construction jobs and 900 permanent high-skilled jobs
	Argentina	OpenAI + Sur Energy	Up to $25bn (LOI, 2025)	Proposed "Stargate Argentina" AI data-center complex (up to 500 MW) powered by advanced nuclear SMR technology	generate thousands of highly skilled jobs in AI, nuclear energy, and construction
MIDDLE EAST	**UAE**	G42 / Khazna + OpenAI, Nvidia, Oracle, Cisco, SoftBank (Stargate UAE)	$Multi-bn	Launch of "Stargate UAE," a 1 GW AI compute cluster forming Phase 1 of a planned 5 GW AI campus	
		Microsoft + G42 / Khazna	$15.2bn (2023–2029; expanded 2024–25)	Major expansion of national AI and cloud capacity, including 200 MW of additional data-center infrastructure, sovereign cloud services, and broader AI service deployment	
	Saudi Arabia	DataVolt + NEOM	$5bn (2025)	Development of a 1.5 GW net-zero AI factory campus in NEOM's Oxagon	Contribute about $320bn to the Middle East economy by 2030
		HUMAIN (PIF-backed) + Nvidia	500 MW (value n/a)	Partnership to build AI factories using Nvidia Blackwell chips, creating large dedicated compute capacity for national AI applications.	
EMERGING EUROPE	**Poland**	Microsoft	$700mn (2025–26)	Second-phase expansion of Poland's hyperscale cloud and AI infrastructure	
	Greece	Data4	$330mn (2024)	a large AI-ready data-center hub to serve Southeast Europe's fast-growing cloud and AI workloads	Creates significant construction and operational employment
	Hungary	ParTec AG	EUR 3bn (2025)	New data-center complex intended to double the capacity of Germany's Jülich supercomputing center	Generates high-skilled jobs in HPC, AI, and data-center operations

Source: Corporate announcements (Google, Microsoft, Amazon Web Services, Nvidia, and OpenAI) and media reports

APPENDIX 4:
AI-related Spending and Economic Impact of Selected Asian Countries (2024)

	CURRENT SPENDING ($ bn)				
	AI		Cloud		ECONOMIC IMPACT (% of GDP)
	Amount	% of GDP	Amount	% of GDP	
China	44.8	0.25	78.4	0.43	-
India	3.3	0.09	17.2	0.44	2.6
Thailand	0.5	0.09	3.3	0.62	1.2
Viet Nam	0.2	0.05	1.4	0.3	0.9
Indonesia	0.4	0.03	2.5	0.18	0.8
Malaysia	0.5	0.11	3.2	0.73	1.3
Philippines	0.3	0.07	2	0.43	1.2
Korea, Rep.	3.4	0.18	9.6	0.51	2.1

Note: AI-related spending includes hardware and software acquisition, fees paid to cloud service providers, and systems integration consultation (IDC "Worldwide AI and generative AI Spending Guide") **Source**: ADB (2025); data from IDC (2024, 2025) and Katz et al. (2025a).

CONTACTS

DEC-Private Markets

- **Dilek Aykut**
  Principal Economist
  daykut@worldbank.org
- **Selin Ozyurt**
  Economist
  sozyurt@ifc.org
- **Kyusam Jung**
  Senior Economist
  kjung1@ifc.org

Infrastructure Chief Economist Office

- **Estefania Vergara Cobos**
  Economist
  evergaracobos@worldbank.org