



The Dialogue®
INFORM ENGAGE IDEATE



RESEARCH REPORT

Building India's

Data Centre Ecosystem

A National Policy Framework for
Sustainable and Sovereign
Digital Infrastructure



SEPTEMBER 2026



RESEARCH REPORT

Building India's Data Centre Ecosystem



A National Policy Framework for Sustainable and Sovereign
Digital Infrastructure

WITH SUPPORT FROM

Intel Corporation

Authors: Soham Jagtap, Anuradha Bhattacharya and Kamesh Shekar

Copyeditor: Akriti Jayant

Thematic Designer: Shivam Kulshrestha

The Dialogue is a public policy think tank with a vision to drive a progressive narrative in India's policy discourse. Founded in 2017, we believe in facilitating well-researched policy debates at various levels to help develop a more informed citizenry, on areas around technology and development issues. The Dialogue has been ranked as the world's Top 10 think tanks to watch out for, by the Think Tank and Civil Societies Programme (TTCSP), University of Pennsylvania in their 2020 and 2021 rankings.

For more information
visit thedialogue.org.in

Suggested citation:
Jagtap, S., Bhattacharya, A. & Shekar, K. (2026). *Building India's data centre ecosystem: A national policy framework for sustainable and sovereign digital infrastructure* [Research report]. The Dialogue.

Catalogue no:
TD/PDG/RR/0926/12

Publication date:
September 16, 2026

Disclaimer
The facts and information in this report may be reproduced only after giving due attribution to the authors and The Dialogue.

Foreword

Architecting Sovereign Compute for the Intelligence Age



We are living through a decisive inflection point in global technological power. Artificial Intelligence is a civilization-transforming technological innovation – a bedrock of national competitiveness, economic catalyst, and strategic autonomy. At the heart of this shift lies physical AI infrastructure: the data centres that house, power, and execute the world's intelligence.

As nations race to secure their digital futures, compute infrastructure can no longer be treated as passive, commercial real estate. It is a sovereign strategic infrastructure – as foundational to national security and economic self-reliance in the 21st century as power grids, trade corridors, and energy reserves were in the 20th.

India sits at an extraordinary juncture. Consuming nearly a fifth of the world's data yet hosting a fraction of global compute capacity, the nation faces both a structural deficit and a generational opportunity. As the AI paradigm transitions from frontier model training to widespread, real-time inference, the demand for resilient, localized compute will explode. Yet, true sovereignty cannot be achieved by simply scaling gigawatts or procuring homogeneous hardware clusters.

Instead it requires a multi-dimensional approach:

- **Strategic Autonomy & Heterogeneity:** Moving beyond reliance on single-vendor hardware monopolies, vertically-integrated AI platform stacks to build inference-ready, heterogeneous compute architectures – blending CPUs, GPUs, NPUs, and custom silicon; providing model selection open vs. frontier; orchestrating compute environments between local, edge and cloud – that ensure choice, resilience, and lock-in avoidance.
- **Distributed Architecture:** Dispersing infrastructure beyond traditional metropolitan clusters into Tier-2 and Tier-3 regional hubs, placing latency-sensitive inference capacity closer to users while protecting the national grid from localized demand shocks.
- **Sustainability by Design:** Integrating clean energy, advanced liquid cooling, water conservation, and waste-heat recovery directly into the operational DNA of the infrastructure, proving that sovereign compute expansion and net-zero commitments can advance together.
- **Community-Centric Local Impact:** Embedding transparent and inclusive engagement with local communities into facility planning – ensuring that data center developments deliver direct environmental stewardship, localized energy grid upgrades, and direct economic and workforce benefits to the surrounding communities.
- **Open Interoperability & Democratized Access:** Ensuring software portability across open standards so that startups, public institutions, and domestic enterprises can affordably access compute without restrictive re-engineering.

Through Aikam – Telangana's state-backed autonomous sovereign AI initiative – Telangana is already translating these core principles into direct action. By adopting a pragmatic, phased investment roadmap, building sovereign compute infrastructure, establishing robust local talent pipelines, and mobilizing global public-private partnerships, Telangana is actively transforming the AI trajectory in Telangana. The model demonstrates that when a region aligns sovereign energy grids with open, heterogeneous compute and cross-border innovation corridors, it can rapidly establish itself as a globally competitive AI powerhouse while preserving technological self-reliance.

This research report, *Building India's Data Centre Ecosystem: A National Policy Framework for Sustainable and Sovereign Digital Infrastructure*, provides a timely, rigorous, and well-crafted synthesis of these imperative themes. It moves past surface-level investment promotion to offer policymakers, industry leaders, and builders a practical, phased blueprint for action.

I invite policymakers, enterprise leaders, investors, and technologists not merely to read this report as an analytical synthesis, but to adopt its principles as a catalyst for immediate policy development and execution. By building the right compute, in the right places, powered by sustainable energy and open architectures, India can transform the coming inference surge into a dominant engine of sovereign digital power.

Phani Nagarjuna

Sovereign AI Strategist, Founding Global CEO of Aikam,
Silicon Valley Executive & Entrepreneur

Table of Contents

ACKNOWLEDGEMENTS	I
LIST OF ABBREVIATIONS	II
EXECUTIVE SUMMARY	IV
1. INTRODUCTION: DATA CENTRES AS STRATEGIC INFRASTRUCTURE	1
2. INDIA'S DATA CENTRE LANDSCAPE	3
2.1 Demand Drivers	3
2.2 Why India Is an Attractive Location	5
2.3 The Policy and Regulatory Landscape	6
2.4 Concerns to Acknowledge and Address	7
3. EXISTING POLICY LANDSCAPE AND GAPS	12
3.1 Existing Policies in India	12
3.2 International Policy Landscape	13
3.3 Key Policy Gaps in India	15
4. GUIDING PRINCIPLES FOR A NATIONAL DATA CENTRE POLICY	18
4.1 Strategic Autonomy	18
4.2 Heterogeneous, Inference-Ready Compute	18
4.3 Sustainability by Design	19
4.4 Distributed Infrastructure	20
4.5 Interoperability and Open Standards	21
4.6 Access, Affordability, and Ecosystem Development	21
4.7 Skill & Workforce Development	22
5. KEY RECOMMENDATIONS AND IMPLEMENTATION ROADMAP	24
5.1 Governance and Regulatory Alignment	24
5.2 Infrastructure and Spatial Planning	24
5.3 Compute Architecture and Access	24
5.4 Sustainability and Energy Integration	25
5.5 Workforce and Ecosystem Development	25
5.6 Implementation Roadmap	25
6. CONCLUSION	27

Acknowledgements

This study reflects the collective insights and contributions of 50 stakeholders from government, industry, academia, financial institutions, technology companies, start-ups and civil society. We are deeply grateful to them for sharing their time, expertise and perspectives, which were instrumental in shaping and refining the study.

We are particularly grateful to Mr Kamesh Shekar and Mr Kazim Rizvi for their guidance, support and encouragement throughout the research. We also thank Ms Akriti Jayant for her support in finalising the study, Ms Twinkle for her editorial support and Mr Shivam Kulshrestha for designing the report. Finally, we sincerely thank all individuals who contributed to this study, whether named or unnamed. Any errors or omissions remain solely our responsibility.

The acknowledgement of any individual or institution does not imply their endorsement of the study's findings, analysis, conclusions or recommendations. The views and interpretations expressed in this report should not be attributed to any individual or institution acknowledged herein.

List of Abbreviations

Abbreviation	Full Form
AI	Artificial Intelligence
AI/ML	Artificial Intelligence / Machine Learning
APAC	Asia-Pacific
ASIC	Application-Specific Integrated Circuit
CAGR	Compound Annual Growth Rate
CHP	Combined Heat and Power
CNI	Critical National Infrastructure
CPU	Central Processing Unit
CTS	Craftsmen Training Scheme
DCEZ	Data Centre Economic Zone
DCFU	Data Centre Facilitation Unit
DCIC	Data Centre Industry Council
DCIS	Data Centre Incentivisation Scheme
DPI	Digital Public Infrastructure
DPDP Act	Digital Personal Data Protection Act, 2023
EED	Energy Efficiency Directive
FPGA	Field-Programmable Gate Array
GCC	Global Capability Centre
GDPR	General Data Protection Regulation
GPU	Graphics Processing Unit
GW	Gigawatt
HVAC	Heating, Ventilation and Air Conditioning
IMEC	Inter-Ministerial Empowered Committee
IT	Information Technology
ITI	Industrial Training Institute
IT-ITeS	Information Technology and Information Technology-enabled Services
LLMOps	Large Language Model Operations
MeitY	Ministry of Electronics and Information Technology
MSME	Micro, Small and Medium Enterprise
MW	Megawatt
NAPS	National Apprenticeship Promotion Scheme
NIS	Network and Information Systems
NIS2	Network and Information Systems Directive 2

NITI Aayog	National Institution for Transforming India
NPU	Neural Processing Unit
NSDC	National Skill Development Corporation
NSM	National Supercomputing Mission
PLI	Production-Linked Incentive
PMKVY	Pradhan Mantri Kaushal Vikas Yojana
PPA	Power Purchase Agreement
PUE	Power Usage Effectiveness
SMR	Steam Methane Reforming
TPU	Tensor Processing Unit
TRAI	Telecom Regulatory Authority of India
UK	United Kingdom
US	United States
WUE	Water Usage Effectiveness

Executive Summary

India's data centre ecosystem is entering a decisive phase. The rapid growth of artificial intelligence (AI), cloud services, digital public infrastructure (DPI) and other data-intensive industries is transforming data centres from commercial facilities into strategic infrastructure. India combines a large digital user base and expanding AI adoption with relatively low existing capacity, creating a significant opportunity to strengthen both domestic and global compute infrastructure. Realising this opportunity, however, will require more than simply adding capacity. It will require India to build the right data centres, in the right locations, with the right compute architectures, energy systems and supporting ecosystems.

India's emerging infrastructure must reflect the shift towards recurring and geographically dispersed AI inference demand. This requires heterogeneous compute architectures combining CPUs, Graphics Processing Unit (GPUs), Neural Processing Unit (NPUs), Tensor Processing Units (TPUs), FPGAs and Application-Specific Integrated Circuits (ASICs) according to workload requirements, rather than treating AI readiness as synonymous with GPU deployment. Large hyperscale facilities should be concentrated in locations with adequate demand and infrastructure, while distributed edge facilities can support latency-sensitive applications and regional digital ecosystems in Tier 2 and Tier 3 cities.

The present policy landscape remains fragmented across central regulations and state-level incentives. Data centre capacity is also concentrated in a few metropolitan corridors, increasing pressure on electricity, water and connectivity infrastructure. Without coordinated planning, rapid expansion could intensify grid constraints, regional imbalances and dependence on imported hardware and proprietary technology stacks.

Developed through extensive stakeholder consultations and secondary research, this paper proposes a National Data Centre Policy built around strategic autonomy, heterogeneous and inference-ready compute, sustainability by design, distributed infrastructure, interoperability, affordable access and workforce development. It recommends a Centre-led framework implemented through empowered state institutions with the Centre defining common standards and national priorities while states retain responsibility for locally appropriate execution and incentives. A federated single-window clearance system should connect national coordination with state-level implementation rather than centralising local approvals. The framework should incentivise regional and edge infrastructure in Tier 2 and Tier 3 cities, while aligning the location and architecture of compute infrastructure with local demand and infrastructure readiness. It should promote heterogeneous, inference-ready compute, alongside interoperability and open standards that enable workloads to move across hardware and cloud ecosystems without restrictive re-engineering.

Sustainability must be embedded into this expansion by integrating Power Usage Effectiveness (PUE) and Water Usage Effectiveness (WUE) standards, linking fiscal incentives to renewable energy integration and resource optimisation, and developing phased waste-heat recovery requirements for large facilities. The policy should simultaneously strengthen the domestic ecosystem through incentives for data centre equipment, semiconductor-linked supply chains and advanced cooling technologies.

The objective should be to develop a coherent national compute ecosystem that is resilient, accessible, globally competitive and aligned with India's energy, sustainability and technological ambitions. Implementation should follow a phased approach, beginning with national policy and institutional coordination, followed by regional infrastructure, sustainability and skilling measures, and ultimately progressing towards interoperability, renewable-energy requirements, affordable compute access and stronger domestic supply chains.

India Data Centre Ecosystem at a Glance



CURRENT CAPACITY

~1.57 GW

India's operational data centre capacity is estimated at around 1.57 GW.



PROJECTED 2030 CAPACITY

8 GW

Capacity is expected to reach around 8 GW by 2030, with significantly higher compute demand driven at AI, cloud and digital public infrastructure.



GEOGRAPHIC CONCENTRATION

~70%

Nearly 70% of India's data centre capacity is concentrated in the Mumbai and Chennai corridors.



INDIAAI COMPUTE ACCESS

~18,000 GPUs

The IndiaAI Compute Access initiative is providing access to around 18,000 GPUs to accelerate domestic AI innovation.



INVESTMENT PIPELINE

~INR 3-5 lakh crore

Announced and expected investments in India's data centre ecosystem are estimated at around INR 3-5 lakh crore over the coming years.



WORKFORCE

~443,000

India's data centre and related digital infrastructure sectors could generate around 443,000 direct and indirect jobs by 2030.

A MORE CONNECTED, SUSTAINABLE AND SOVEREIGN DIGITAL FUTURE

India's data centre ecosystem is at an inflection point. With the right policies, infrastructure, energy systems and skills, it can power an inclusive, innovative and resilient digital economy.

1. Introduction: Data Centres as Strategic Infrastructure

Data centres have emerged as one of the most critical foundations of the digital economy. Once viewed primarily as facilities for storing and processing data, they have evolved into strategic infrastructure that underpins Artificial Intelligence (AI), cloud computing, Digital Public Infrastructure (DPI), e-commerce, financial services, telecommunications, scientific research, and an increasingly digitised public sector. As governments and businesses rely more heavily on digital technologies, the availability of secure, resilient, and high-performance data centre infrastructure has become indispensable to economic growth, innovation, and national competitiveness. Data centres should not be understood merely as digital infrastructure, but as strategic national assets whose significance extends beyond commercial hosting to encompass economic resilience, national security, and geopolitical competitiveness. Much like ports, power grids, and telecommunications networks, they constitute foundational infrastructure upon which a country's digital economy and critical public services depend.

The rapid advancement of AI has further transformed the strategic importance of data centres. Unlike earlier generations of digital services, AI systems require substantial compute capacity, increasingly delivered through specialised accelerators such as Graphics Processing Units (GPUs), Neural Processing Units (NPUs), Tensor Processing Units (TPUs), Field-Programmable Gate Arrays (FPGAs), and Application-Specific Integrated Circuits (ASICs). In tandem, demand is growing for both traditional and AI-accelerated Central Processing Units (CPUs), which support data preprocessing, workload orchestration, coordination, and system management. These systems also depend on low-latency connectivity, reliable power supplies, and sufficient cooling infrastructure. As AI models continue to scale and become more complex, demand for compute infrastructure is expected to grow significantly, placing data centres at the centre of discussions on digital infrastructure, energy systems, industrial policy, and technological sovereignty. Consequently, data centres are no longer passive storage facilities; they have become the physical backbone of the AI economy.

At the same time, data centres increasingly intersect with national priorities extending beyond the digital sector. For India, the importance of data centres is amplified by the country's rapidly expanding digital ecosystem. India has built one of the world's largest DPI ecosystems, accelerated digital adoption across sectors, and emerged as one of the fastest-growing AI markets globally. Government initiatives such as IndiaAI, increasing cloud adoption, the growth of digital public services, and the expansion of AI-enabled industries are expected to significantly increase demand for data storage and compute infrastructure. Simultaneously, India is witnessing substantial investments in hyperscale data centres, edge computing, cloud infrastructure, and submarine cable connectivity. These developments present a unique opportunity for India to position itself as a global hub for digital infrastructure and AI-ready compute capacity.

However, this rapid growth also presents a new set of policy challenges. Data centre development has implications for electricity demand, renewable energy integration, water consumption, land use, digital connectivity, cybersecurity, supply chains, and regional development. Uncoordinated expansion risks creating infrastructure bottlenecks, resource constraints, and regional imbalances that could undermine both sustainability and long-term competitiveness. Moreover, as AI increasingly becomes a strategic technology, policy discussions must move beyond promoting physical infrastructure towards building sovereign compute capacity that supports innovation, economic resilience, and trusted digital ecosystems.

While India has introduced several policy initiatives to encourage investment in data centres, including the Draft Data Centre Policy, 2020 and various state-level policies, the existing landscape remains fragmented. Current approaches primarily focus on facilitating investment and improving ease of doing business, with comparatively limited attention to issues such as AI-driven compute demand, integrated energy planning, sustainability, cybersecurity, equitable access to compute resources, and long-term governance. As data centres evolve from digital infrastructure to strategic national assets, there is a growing need for a comprehensive policy framework that aligns infrastructure development with India's broader digital, industrial, energy, and AI ambitions.

Against this backdrop, this paper argues that India requires a National Data Centre Policy that adopts a holistic and forward-looking approach to governing the country's data centre ecosystem. Such a policy should move beyond investment promotion to address the full spectrum of challenges associated with data centre development, including spatial planning, energy integration, sustainability, compute architecture, connectivity, cybersecurity, ecosystem development, workforce capacity, and institutional coordination. By doing so, India can transition from simply expanding digital infrastructure to building sustainable and sovereign compute infrastructure that supports AI innovation, strengthens strategic autonomy, and positions the country as a globally competitive destination for the next generation of digital and AI technologies.

The paper begins by examining the evolution of India's data centre landscape and the key drivers shaping its growth. It then analyses the existing policy framework and identifies the principal regulatory and institutional gaps that constrain the development of a coordinated national strategy. Building on this analysis, the paper proposes a set of guiding principles and outlines the core pillars of a National Data Centre Policy, followed by an implementation roadmap to support its phased adoption. In doing so, the paper seeks to provide a comprehensive policy framework for developing a resilient, sustainable, secure, and globally competitive data centre ecosystem that advances India's digital economy while supporting its long-term AI and strategic ambitions.

2. India's Data Centre Landscape

In June 2026, Meta and Reliance Industries announced a partnership to build an AI-enabled data centre of 168 megawatts (MW) at Jamnagar, Gujarat, with Reliance designing, constructing and operating the facility and Meta leasing its capacity. The plant, which will be powered entirely by renewable energy and cooled with desalinated seawater, sits within a far larger campus that Reliance has designated for up to three gigawatts (GW) of future capacity.¹ This announcement arrived alongside comparable commitments across several states, including the AdaniConneX and Google campus at Visakhapatnam in Andhra Pradesh, reported at an investment of around 15 billion US (United States) dollars over five years.² The logic behind this build-out is the rising demand for AI-ready infrastructure, cloud computing and the substrate on which emerging technologies such as edge computing depend. Data localisation requirements under sectoral regulations, together with the compute ambitions of the IndiaAI Mission, lend the expansion a regulatory and policy momentum of its own, since the graphics processing unit clusters the state is procuring will require co-location infrastructure at scale. Based on current projections, India's data centre capacity is expected to reach between five and eight GW by 2030, more than tripling and, on some estimates, rising fourfold from a present base of roughly 1.4 to 1.6 GW, supported by investments in the range of 30 to 35 billion US dollars.³

2.1 Demand Drivers

Globally, the structural foundation of this growth is driven by the surge in energy demand for AI training and inference. Inference models, in particular, meaning AI models that make decisions or generate outputs in real-time using real-world data are expected to dominate the share of data centre energy demand over the next five years.

As is discernible from the graph below, energy demand from data centres is expected to grow sharply over the next five years, with total energy demand more than doubling from the current 82.3 GW to 219 GW in 2030, representing a compound annual growth rate (CAGR) of 22%. Interestingly, data centre energy demand for AI inference is expected to triple from the current 31.2 GW to 93.3 GW in 2030, with its share of total demand catapulting to 42.6%, overtaking both non-AI and AI training workloads. To account for the surge in energy demand from data centres, particularly for AI inference models, it is necessary to have an infrastructural, policy and investment strategy to support the build-out of more data centres.

For India, this situation presents a golden opportunity to become an influential player in the global AI supply chain by providing the infrastructure (data centres and compute) to power AI inferencing in addition to AI training and conventional workloads. The strategy must be balanced to ensure that we leverage our capabilities to provide inference-ready, heterogeneous compute to gain a stronger foothold in the AI supply chain and benefit geopolitically. The build-out must also be evaluated against sustainable practices to ensure India's sustainable development goals are not ignored during this process.

¹ NITI Frontier Tech Hub. (2026, February). *India's technology services: Reimagination ahead*. NITI Aayog, Government of India. <https://www.niti.gov.in/node/2184>

² Jain, M. A. (2025, December). *Democratising access to AI infrastructure* (White Paper Series version 3.0). Office of the Principal Scientific Adviser to the Government of India. https://psa.gov.in/CMS/web/sites/default/files/publication/WP_Democratising%20Access_V3.0_29122025A.pdf

³ Sen, A. (2025). *Building India's data centres* (Takshashila Report 2025-27). The Takshashila Institution. <https://takshashila.org.in/research/building-indias-data-centres>



Figure 1: Global Data Centre Demand Workload, 2025-30⁴

India's near-term opportunity is increasingly tied to AI deployment rather than frontier-model training. Its large user base, expanding enterprise adoption, multilingual applications and public-sector use cases are creating demand for inference, sovereign compute and low-latency local deployment.⁵ This shift changes the composition of infrastructure required. Frontier-model training is typically GPU-dominant, whereas production inference, particularly agentic AI, requires greater CPU capacity for orchestration, data retrieval, tool calls, memory management and other general-purpose tasks alongside accelerator-based model execution. Intel has indicated that the CPU-to-GPU ratio of approximately 1:8 associated with frontier training may move towards 1:1, or greater CPU density, for agentic AI. Market analysis similarly reports movement from 1:8 to 1:4 in active agentic deployments, with a trajectory towards parity.^{6,7} These ratios should be treated as directional rather than as a universal design specification, since the optimal configuration will vary according to the model, workload and deployment environment. Nevertheless, they demonstrate why India's infrastructure strategy should not equate AI readiness with GPU procurement alone. It should support balanced CPU-GPU configurations alongside NPUs, ASICs, memory, networking and storage suited to inference-intensive workloads.

This shift towards deployment and inference strengthens the domestic case for locally distributed data centres built around heterogeneous compute. Locating such capacity closer to demand can reduce latency, improve responsiveness and allow infrastructure to be configured around regional workloads. Data centres can combine general-purpose CPUs, GPUs and purpose-built AI accelerators in configurations tailored to particular training and inference requirements.^{8,9} The appropriate hardware mix will vary according to the model, workload and deployment environment. Matching workloads to suitable compute architectures can

⁴ McKinsey & Company. (2025). *Global data centre demand by workload, 2025-30*. McKinsey & Company.

<https://www.mckinsey.com/featured-insights/charts/data-center-demands>

⁵ New Market Pitch Team. (2026, June 29). *Why is India attracting more AI infrastructure lately?* New Market Pitch.

<https://newmarketpitch.com/blogs/news/ai-infrastructure-india>

⁶ MarketsandMarkets. (n.d.). *AI inference CPU market size, share & growth report 2025-2032*. Retrieved August 27, 2026, from <https://www.marketsandmarkets.com/Market-Reports/ai-inference-cpu-market-233319695.html>

⁷ Intel. (2026, June 2). *Computex 2026: An intelligent world built on silicon*. Intel Newsroom.

<https://newsroom.intel.com/artificial-intelligence/computex-2026-an-intelligent-world-built-on-silicon>

⁸ International Energy Agency. (2025). *Energy demand from AI*. In *Energy and AI*. <https://www.iea.org/reports/energy-and-ai/energy-demand-from-ai>

⁹ MarketsandMarkets. (n.d.). *AI inference CPU market size, share & growth report 2025-2032*. Retrieved August 27, 2026, from <https://www.marketsandmarkets.com/Market-Reports/ai-inference-cpu-market-233319695.html>

improve hardware utilisation and resource efficiency, without assuming that any single processor category is appropriate for every use case.

A heterogeneous build-out of compute infrastructure can help align compute capacity and electricity demand with the workload profile of a region. For example, a cluster of hospitals could use locally hosted, domain-specific language-model services delivered from a nearby regional or edge data centre. For modest, latency-sensitive or low-batch inference workloads, servers based on CPUs with integrated vector and matrix acceleration, such as AVX-512 VNNI and Intel Advanced Matrix Extensions, may be suitable. Larger models or high-concurrency workloads may instead require GPUs or purpose-built AI accelerators, including NPUs where independently benchmarked performance justifies their use.¹⁰ Capacity should therefore be provisioned against measured model size, numerical precision, memory capacity and bandwidth, latency, concurrency and throughput requirements, rather than through a predetermined number of NPU racks. Matching hardware to the workload can improve utilisation and energy per inference, but no compute category is inherently more energy-efficient across all workloads.¹¹ Regional and edge deployment can also reduce latency and wide-area data movement and support compliance with applicable data-residency, storage and governance requirements.¹² This should not, however, be framed as a blanket data-localisation obligation as Section 16 of India's Digital Personal Data Protection Act, 2023 (DPDP Act) permits cross-border transfers subject to government restrictions while preserving stricter sector-specific rules.

2.2 Why India Is an Attractive Location

The clearest argument for investment is the gap between consumption and capacity. As discussed in the previous chapter, the energy for inference models is expected to triple over the next four years. India is estimated to account for roughly one-fifth of global data consumption while hosting only around 5–6 per cent of global data-centre capacity.¹³ Separate industry estimates place India's data-centre capacity density at approximately 1 MW per million internet users, compared with around 51 MW in the United States and 4 MW in China.¹⁴ Low penetration, a surge in AI inference workloads, and India's growing internet user base necessitate structural catch-up.¹⁵

Relatively lower construction costs make India a lucrative destination for investors and hyperscalers. Construction costs in India average around six to seven million US dollars per MW, materially below mature markets such as Singapore and Japan¹⁶ and these are complemented by competitive power tariffs, a large pool of technical talent and favourable submarine cable geography, with major landing stations clustered around Mumbai and Chennai. The surge in demand for data centres has, in turn, increased the demand for skilled blue-collar workers, particularly electricians and Heating, Ventilation, and Air Conditioning (HVAC) technicians. Hyperscaler operators are aware of the surge in demand and are working constantly to shore up the future pipelines of skilled blue-collar workers. Recently, Meta launched a workforce academy in the US, a five-week training programme with a job guarantee at the end of it to counteract the dearth of skilled blue-collar workers in the US.¹⁷

The fiscal architecture in India has also matured in parallel, benefiting the sector. Instruments such as single-window clearances at the state level, extended tax exemptions and Goods and Services Tax (GST)

¹⁰ Intel Corporation. (2024, February 14). *Accelerate PyTorch training and inference using Intel AMX*. <https://www.intel.com/content/www/us/en/developer/articles/technical/accelerate-pytorch-training-inference-on-amx.html>

¹¹ Reuther, A., Michaleas, P., Jones, M., Gadepally, V., Samsi, S., & Kepner, J. (2022). AI and ML accelerator survey and trends. In *2022 IEEE High Performance Extreme Computing Conference (HPEC)* (pp. 1–10). IEEE. <https://doi.org/10.1109/HPEC55821.2022.9926331>

¹² Shi, W., Cao, J., Zhang, Q., Li, Y., & Xu, L. (2016). Edge computing: Vision and challenges. *IEEE Internet of Things Journal*, 3(5), 637–646. <https://doi.org/10.1109/JIOT.2016.2579198>

¹³ Dugal, I. (2025, December 3). India file: \$100 billion data centre boom tests resource limits. *Reuters*. <https://www.reuters.com/world/india/india-file-100-billion-data-centre-boom-tests-resource-limits-2025-12-03/>

¹⁴ Financial Times. (2025, May 12). Low-cost India seen as potential regional hub in data centre boom. <https://www.ft.com/content/7f4aff43-b181-4965-96bc-420d502673ad>

¹⁵ Deshmukh, I. (2026, May 2). *AI's next trajectory: The inference innovation*. Observer Research Foundation. <https://www.orfonline.org/expert-speak/ai-s-next-trajectory-the-inference-innovati>

¹⁶ Pomnar, A., Rajawat, A. S., Tatkar, N. S., & Bhaladhare, P. (2023). Sustainable power prediction and demand for hyperscale datacenters in India. *Engineering Proceedings*, 59(1), 124. https://doi.org/10.3390/engproc202305912a_4

¹⁷ Meagher, H., & Dhirani, L. L. (2024). Cyber-resilience, principles, and practices. In K. N. Qureshi et al. (Eds.), *Cybersecurity vigilance and security engineering of Internet of Everything* (pp. 57–78). Springer.

benefits in various states have provided much-needed regulatory streamlining for the construction of data centres in India.^{18,19} The Asia-Pacific (APAC) region is projected to attract investments of nearly \$800 billion by 2030. India is expected to capture a significant share of these investments, positioning itself as a major player in the region. To achieve this scale, intention must be carefully transitioned into commissioned capacity.²⁰

2.3 The Policy and Regulatory Landscape

A foundational enabler arrived when data centres exceeding capacity of 5 MW were granted infrastructure status by the Central Government, which eased access to long-tenor project finance on more favourable terms. (Ministry of Finance, 2022)²¹ The more visible driver of data centre location decisions today, however, is competition between states. Maharashtra, Tamil Nadu, Telangana, Karnataka, Gujarat, Uttar Pradesh, Haryana, Odisha and West Bengal have moved beyond broad digital policies to dedicated data centre frameworks, each offering a distinct package of incentives.²² Maharashtra has offered substantial electricity duty exemptions over 15 years, along with stamp duty relief and assured round-the-clock supply through state distribution companies. Telangana has provided full reimbursement of stamp duty and registration fees. Tamil Nadu has emphasised dedicated power feeders, concessional tariffs and land subsidies, and Uttar Pradesh has combined capital subsidies with transmission and wheeling charge exemptions.^{23,24} This allows states to tailor offerings according to their endowments and to compete on speed to market.

The same diversity, however, produces a fragmented landscape that the industry has aptly characterised as one nation operating under many laws.²⁵ Each state functions, in effect, as a separate regulatory jurisdiction, requiring distinct legal strategies and localised advisers, which adds overhead, increases capital expenditure, and delays market entry. Single-window clearance, though widely advertised, often remains aspirational in practice, with operators reporting the need for thirty or more approvals across departments and tiers of government.²⁶ The absence of a binding national framework compounds the problem. A Draft National Data Centre Policy was circulated in 2020 but was never notified,²⁷ leaving the Centre's role largely to incentive signalling rather than harmonisation. The Telecom Regulatory Authority of India (TRAI) has sought to fill part of this vacuum, recommending a National Single Window System for time-bound clearances as part of a broader effort to streamline the data economy.²⁸

¹⁸ Government of Maharashtra. (2023). *Maharashtra IT-ITeS policy 2023*. Industries Department. https://industry.maharashtra.gov.in/sites/default/files/2025-09/it-policy-booklet_1_11zon.pdf

¹⁹ The concern herein lies with the fragmentation of regulatory and licensing policies across states, as each state boasts a unique dedicated data centre or IT/ITES policy tailored to its market conditions. The formalisation of a national data centre policy shall go a long way in streamlining approval processes and licensing requirements.

²⁰ Shafaghat, A. (2025). A fuzzy-set qualitative comparative analysis (fsQCA) of dynamic capabilities and microfoundations for hyperscale data center project delivery. *Buildings*, 16(11), 2284. <https://doi.org/10.3390/buildings16112284>

²¹ Ministry of Finance. (2022, October 11). *Notification: Amendment to the Harmonized Master List of Infrastructure Sub-sectors* (Gazette of India, Extraordinary, Part II, Section 3, Sub-section (ii)). Department of Economic Affairs, Government of India. <https://dea.gov.in/>

²² KPMG International. (2026, July). *India data centre opportunity: From emerging demand hub to integrated data centre powerhouse*. KPMG Insights. <https://kpmg.com/in/en/insights/2026/07/india-data-centre-opportunity-from-emerging-demand-hub-to-integrated-data-centre-powerhouse.html>

²³ Government of Tamil Nadu. (2021). *Tamil Nadu data centre policy 2021* (G.O. Ms. No. 33). Information Technology Department, Government of Tamil Nadu. <https://spc.tn.gov.in/policy/tamil-nadu-data-centre-policy-2021/>

²⁴ Government of Uttar Pradesh. (2022). *Uttar Pradesh data centre policy 2021 (Amended 2022)* (Section 4.2). Invest UP, Department of IT & Electronics. <https://invest.up.gov.in/up-data-centre-policy-2021/>

²⁵ Anarock, & Mace. (2023). *Sustaining the momentum: India's data centre growth story*. Anarock Property Consultants. <https://websitemedia.anarock.com/>

²⁶ Meagher, H., & Dhirani, L. L. (2024). Cyber-resilience, principles, and practices. In K. N. Qureshi, T. Newe, G. Jeon, & A. Chehri (Eds.), *Cybersecurity vigilance and security engineering of Internet of Everything* (pp. 57–78). Springer. https://doi.org/10.1007/978-3-031-45162-1_3

²⁷ Meagher, H., & Dhirani, L. L. (2024). Cyber-resilience, principles, and practices. In K. N. Qureshi, T. Newe, G. Jeon, & A. Chehri (Eds.), *Cybersecurity vigilance and security engineering of Internet of Everything* (pp. 57–78). Springer. https://doi.org/10.1007/978-3-031-45162-1_3

²⁸ Telecom Regulatory Authority of India. (2022, November 18). *Recommendations on regulatory framework for promoting data economy through establishment of data centres, content delivery networks, and interconnect exchanges in India*. TRAI, Government of India. https://traigov.in/sites/default/files/2024-09/Recommendations_18112022.pdf

The practical cost of fragmentation is measured in time. Owing to the layering of state and central requirements, the path from land acquisition through the resolution of title issues to the receipt of necessary permits is long and uncertain. As a result, construction that should ordinarily take around a year can extend to three or four years in India, and even in the better-organised states with designated data centre parks the timeline from land deal to commercial operation runs to roughly eighteen to twenty-four months before any complications.²⁹ Elongated timelines deter investment and strain the economics of partnerships, which is why a streamlined, coordinated approval process is among the most consequential reforms available to policymakers. Streamlining should not be equated with transferring project-level approvals to the Centre. A more workable model is a nationally consistent framework implemented through empowered state institutions, with responsibility allocated to the level of government best placed to discharge it.

2.4 Concerns to Acknowledge and Address

A clear-eyed account of the landscape must register the tensions that accompany rapid build-out.

2.4.1 Geographic Concentration

Geographic distribution is simultaneously a technical, economic, and governance question. Currently, data centres in India are heavily clustered (refer to the aforementioned illustration). The Mumbai Metropolitan Region and Chennai together account for around seventy per cent of data centres in India, with the bulk of recent capital deployment directed at these two markets.³⁰ This has to do with economics rather than policy. Hyperscalers prefer to locate cloud regions where they can efficiently serve enterprise customers based on factors such as enterprise demand, population density, network connectivity, existing digital ecosystems and latency requirements. Since metropolises with established digital infrastructure are where enterprises house their major operations, demand for compute power and low latency is highest in these regions. Consequently, hyperscalers tend to establish cloud regions in these locations. Access to submarine cable landing stations, cooling infrastructure (water availability), and an uninterrupted power supply are also crucial factors contributing to the concentration of cloud regions in a handful of metropolitan regions. For instance, the Google AI Hub, a 1 GW) AI data centre campus that is planned to be built by 2030, chose Vishakhapatnam as its location because of its access to submarine cable landing stations, large tracts of land, and the ability to build a dedicated high-capacity power infrastructure.³¹

Concentration of this order raises questions of grid loading, regional resilience and balanced development, since a shock to power or connectivity in one corridor carries disproportionate national consequences. Furthermore, the concentration of cloud regions may also impose detrimental strain on potable water resources. Therefore, there is a need for a long-term vision in which computing power is optimally distributed across the nation's geography and tailored to local needs. An encouraging counter-trend is the migration of edge facilities to Tier 2 cities such as Jaipur, Ahmedabad, Visakhapatnam, Kochi, Lucknow, and Patna, driven by the low-latency requirements of real-time AI applications, which over time should distribute both demand and risk more evenly.³² Tier 2 and Tier 3 cities do not yet generate sufficient enterprise demand to justify large GPU clusters, and enterprise and AI workloads remain concentrated in major metropolitan centres. Furthermore, building hyperscale data centres in these locations will pose a significant challenge, given the need for infrastructure and ecosystem readiness, including proximity to subsea cable landing stations, reliable international fibre backbones, and high-capacity utilities. They would benefit more from the build-out of distributed edge infrastructure to meet the expanding demand for consumer services that require content delivery closer to users. Furthermore, as discussed earlier, the demand for inference is likely to increase in the near future, implying that Tier 2 and Tier 3 cities would

²⁹ CBRE India. (2024). *2024 India data centre market update*. CBRE Research.

<https://www.cbre.co.in/insights/reports/2024-india-data-centre-market-update>

³⁰ TOI Business Desk. (2025, November 8). Data centre boom: India's capacity to hit 8GW by 2030; \$30 billion investment seen driving AI, cloud growth. *The Times of India*. <https://timesofindia.indiatimes.com/business/india-business/data-centre-boom-indias-capacity-to-hit-8gw-by-2030-30-billion-investment-seen-driving-ai-cloud-growth/articleshow/125183133.cms>

³¹ Kulkarni, M. (2026, April 20). Google, Meta are building huge data centres in Vizag. Nobody told the power grid. *The Ken*. <https://the-ken.com/story/google-meta-are-building-huge-data-centres-in-vizag-nobody-told-the-power-grid/>

³² Motwani, S. (2026, July 21). The infrastructure revolution no one saw coming: Tier II cities lead the edge. *Fortune India*. <https://www.fortuneindia.com/opinion/the-infrastructure-revolution-no-one-saw-coming-tier-ii-cities-lead-the-edge/149465>

benefit from distributed infrastructure with optimised compute architectures tailored to a region's specific demands. The recommendation in this context is to build large-scale, GPU-heavy hyperscale data centres in geographies with established ecosystems and infrastructure readiness, while purpose-specific, distributed edge facilities that support inference-heavy workloads can be spatially deployed to build local data ecosystems. Distributed data centre infrastructure could also enable organisations to scale into new markets and customer bases.³³

Beyond physical infrastructure, cyber resilience must be embedded into spatial planning itself, as distributed edge and Tier 2 and Tier 3 deployments introduce expanded attack surfaces that require location-specific security architectures.³⁴ Finally, the workforce dimension cannot be overlooked. The benefits of distributed digital infrastructure will only materialise equitably if accompanied by parallel investments in skilling and regional capacity building, ensuring that Tier 2 and Tier 3 cities are not just hosts to data centre parks but active participants in the digital economy they enable.³⁵

³³ Flexential. (2025, February 17). *The value of geographically diverse data centers*.

<https://www.flexential.com/resources/blog/value-geographically-diverse-data-centers>

³⁴ Meagher, H., & Dhirani, L. L. (2023). Cyber-resilience, principles, and practices. In K. N. Qureshi, T. Newe, G. Jeon, & A. Chehri (Eds.), *Cybersecurity vigilance and security engineering of internet of everything* (pp. 57–74). Springer. https://doi.org/10.1007/978-3-031-45162-1_4

³⁵ Kattumuri, R., & Singh, S. (2020). *Empowering labour force in technology driven era: An Indian context* (Working Paper No. 11). India Observatory, London School of Economics and Political Science. https://sticerd.lse.ac.uk/textonly/india/publications/working_papers/IOWP11-KattumuriSingh.pdf

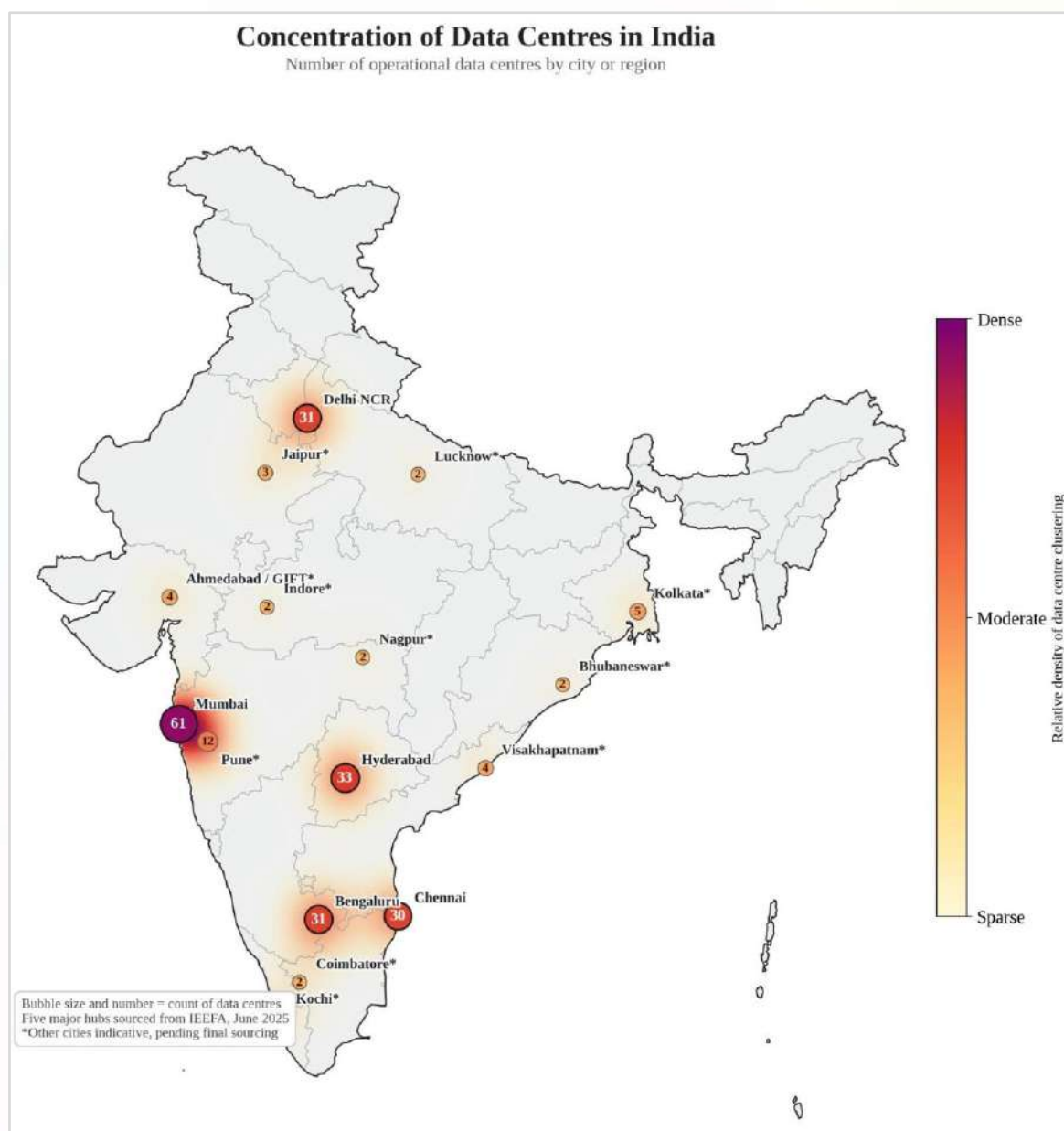


Figure 2: Concentration of Data Centres in India

2.4.2 Power Demand and Grid Readiness

The capacity requirements of the Indian data centre sector are projected to increase from 1.57 GW in August 2026 to 8 GW by 2030³⁶, requiring investments totalling approximately \$300 billion.³⁷ Crucially, the operational framework of these facilities requires uninterrupted, multi-source power distribution. Data centres are inherently energy-intensive facilities, exhibiting energy intensities 10 to 100 times greater than

³⁶ Press Information Bureau. (2026, September 14). *Data centres in India infrastructure for the digital age* (Explainer ID: 159969). Ministry of Information and Broadcasting, Government of India.
<https://pib.gov.in/PressReleasePage.aspx?PRID=2298108>

³⁷ ETDatacenters. (2026, August 26). *India data centre footprint to quadruple by 2030: Anarock*.
<https://datacenters.economictimes.indiatimes.com/news/investments-deals/india-data-centre-footprint-to-quadruple-by-2030-anarock/133528174>

those of standard commercial real estate. This creates a severe demand strain on a domestic grid system that historically operates at narrow margins between peak demand and actual peak power availability.³⁸

The challenge is not limited to generation capacity alone. In several emerging data centre hubs, planned investments are approaching or exceeding the capacity of existing local electricity infrastructure, requiring significant upgrades to transmission networks, substations, and distribution systems. The geographical concentration of hyperscale campuses can create localised demand shocks, raising concerns regarding grid stability and the ability of existing infrastructure to support additional high-density loads.

To build sustainable capacity while advancing renewable energy objectives, industrial operators are executing bilateral Power Purchase Agreements (PPAs) to procure utility-scale solar and wind power. However, integrating these green power sources poses grid-readiness challenges. Renewable energy is inherently intermittent and requires dedicated storage systems to achieve the 24/7 "always-on" reliability mandatory for mission-critical operations. While renewable energy procurement can significantly reduce the carbon footprint of data centres, it cannot, on its own, meet the reliability requirements of mission-critical facilities. Solar generation is variable and dependent on weather conditions, while battery storage technologies capable of providing long-duration, uninterrupted power remain expensive and are still developing at scale. This creates a need for a broader energy strategy combining renewable generation, grid augmentation, energy storage, and reliable backup systems. To mitigate these baseline limitations, advanced facilities utilise AI and ML load-balancing algorithms designed to align data workflows dynamically with real-time renewable energy availability.³⁹ Furthermore, long-term technological alternatives under review include deploying hydrogen gas turbines utilising hydrogen extracted via Steam Methane Reforming (SMR), as well as localised Combined Heat and Power (CHP) tri-generation systems that simultaneously produce electricity, heat, and cooling.

However, technological solutions at the facility level must be complemented by coordinated infrastructure planning at the national and state levels. Future data centre policy should therefore integrate electricity planning with digital infrastructure planning, ensuring that power generation, transmission capacity, renewable energy integration, and data centre expansion are developed in parallel rather than addressed sequentially. The composition of that power matters as much as its quantity. If the build-out leans on coal-fired generation, it sits in direct tension with India's commitment to net-zero emissions by 2070. Scaling renewable capacity in step with data centre demand is therefore not an environmental afterthought but a precondition for the sector's long-term legitimacy and for the credibility of its hyperscale tenants, most of whom carry their own decarbonisation commitments. The recent contracting of close to a GW of clean energy in India by Meta, through CleanMax and Fourth Partner Energy, illustrates how anchor demand can itself pull new renewable capacity onto the grid.⁴⁰

The operational responsibility for aligning power grid readiness with the upcoming data centre build-out should sit primarily with the states. State transmission utilities, distribution companies and regulatory institutions are best placed to map location-specific demand, plan substations and network augmentation, facilitate open access and utility connections, and assess whether proposed clusters are compatible with local grid and water constraints.⁴¹

³⁸ Pomnar, A., Rajawat, A. S., Tatkar, N. S., & Bhaladhare, P. (2023). Sustainable power prediction and demand for hyperscale datacenters in India. *Engineering Proceedings*, 59(1), Article 124. <https://doi.org/10.3390/engproc2023059124>

³⁹ Ganguly, S., Raje, S., Kumar, S., Sartor, D., & Greenberg, S. (2016). *Accelerating energy efficiency in Indian data centers: Final report for phase I activities* (Report No. LBNL-1005085). Lawrence Berkeley National Laboratory. <https://www.osti.gov/biblio/1249186>

⁴⁰ ETSustainability. (2026, June 12). *Meta and CleanMax expand India green energy partnership with major 900 MW boost*. ET Sustainability. <https://sustainability.economictimes.indiatimes.com/news/green-tech/meta-and-cleanmax-expand-india-green-energy-partnership-with-major-900-mw-boost/131674873https://sustainability.economictimes.indiatimes.com/news/green-tech/meta-and-cleanmax-expand-india-green-energy-partnership-with-major-900-mw-boost/131674873>

⁴¹ Government of India. (2003). *The Electricity Act, 2003* (Act No. 36 of 2003, ss. 39, 42 and 86). India Code. https://www.indiacode.nic.in/ViewFileUploaded?file=Electricity+Act+2003.pdf&path=AC_CEN_19_22_00001_200336_1517807317930%2Frulesindividualfile%2F

2.4.3 Managing the Surging Inference Demand

The preceding concerns acquire their urgency from a single structural shift in the nature of AI demand itself. For much of the last decade, the centre of gravity in AI infrastructure was training, a high but essentially one-time cost incurred when a model is built. Inference, by contrast, is recurring, since it is incurred every time a deployed model answers a query, and it therefore grows in step with adoption rather than tapering once development concludes.⁴² The financial consequence is stark. A model that costs in the order of a hundred million US dollars to train can cost more than twice that figure annually to operate at scale.⁴³ This recurring character is what drives the projection, noted earlier in this chapter, that inference will overtake both training and non-AI workloads to dominate global data centre energy demand by 2030.⁴⁴ The market follows the same trajectory, with the global inference infrastructure-as-a-service segment expected to expand from roughly 225 billion US dollars to about 551 billion by 2031.⁴⁵ The energy dimension is equally consequential, as inference, rather than training, accounts for the overwhelming majority of the lifecycle energy a deployed AI system consumes. For India, then, inference is simultaneously the largest commercial opportunity in the compute value chain and the principal source of future grid stress.

India must enter with greater preparedness to capture the opportunity on present terms. Currently, inference-dominant capacity accounts for under a tenth of the country's installed AI compute, with the remainder weighted heavily towards training-optimised silicon procured for an earlier phase of the AI cycle. Much of the prevailing inference toolchain is tightly coupled to a narrow set of proprietary hardware and cloud stacks controlled by a small number of external actors, so that efficiency gains achieved on one vendor's architecture do not port cleanly to alternatives, entrenching dependence even as utilisation improves.⁴⁶ Hence, in this case, efficiency and sovereignty pull apart rather than together. India must strategically place itself to avoid falling into this web. Compounding this, the procurement and empanelment routes through which government and enterprise buyers access compute remain thin on inference-ready options, and the layered approval timelines described above mean that capacity intended to serve real-time workloads is often commissioned years after the demand it was meant to meet has materialised. The result is a country that consumes close to a fifth of the world's data yet hosts a fraction of the inference capacity required by that consumption.

As discussed earlier, India must look to distribute its data centres spatially to encourage the growth of local data ecosystems, increase efficiency, and reduce the disproportionate load on the power grid infrastructure. However, distributed edge facilities cannot be solely GPU-based hyperscalers supporting training workloads, as they require a mature ecosystem and infrastructure readiness. Therefore, distributed edge facilities in tier 2 and tier 3 cities must support heterogeneous compute architectures that blend CPUs, NPU, and specialised application-specific accelerators. This shall not only address the lack of infrastructure and ecosystem readiness, but also support local inference workloads more efficiently whilst reducing reliance on proprietary hardware and software stacks. Policy must encourage open-standard software frameworks and ensure hardware interoperability across different cloud ecosystems. This can ensure that local developers are not tethered to a single, expensive global silicon monopoly.

Meeting this inference demand strategically requires the build-out to be designed around inference from the outset rather than retrofitted to it, and the three threads running through this chapter, heterogeneous compute, sustainability by design, and spatial distribution, are best understood as a single integrated response. Heterogeneity matters because inference workloads are diverse and frequently latency-sensitive, and a facility composed of a mix of GPUs, NPUs, and purpose-built inference accelerators can be matched to the workload profile of the region it serves. This in turn advances the sustainability objective directly, since inference-optimised hardware draws materially less power per unit of useful output, easing the grid burden that an undifferentiated build-out would otherwise impose. Spatial distribution completes the logic, as situating inference capacity in tier 2 and tier 3 cities places compute close to the point of demand, lowers latency for real-time applications, and distributes load across the grid rather than concentrating it in two corridors.

⁴² Deshmukh, I. (2026, May 2). *AI's next trajectory: The inference innovation*. Observer Research Foundation. <https://www.orfonline.org/expert-speak/ai-s-next-trajectory-the-inference-innovation>

⁴³ Gartner. (2025). *Forecast: AI spending, worldwide, 2024–2029 (4Q25)* [as cited in Boston Consulting Group analysis]. Cited in Aikam, Blaize, & Yotta. (2026). *Aikam Cloud AI compute offering*.

⁴⁴ McKinsey & Company. (2025). *Global data centre demand by workload, 2025–30*. McKinsey & Company.

⁴⁵ Mordor Intelligence. (2026). *Infrastructure as a service market size & share analysis - growth trends and forecast (2026 - 2031)*. <https://www.mordorintelligence.com/industry-reports/infrastructure-as-a-service-market>

⁴⁶ Deshmukh, I. (2026, May 2). *AI's next trajectory: The inference innovation*. Observer Research Foundation. <https://www.orfonline.org/expert-speak/ai-s-next-trajectory-the-inference-innovation>

3. Existing Policy Landscape and Gaps

India currently governs data centres through several state-level policies aimed at attracting investment.⁴⁷ While the Draft Data Centre Policy, 2020, released by the Ministry of Electronics and Information Technology (MeitY), sought to establish a comprehensive framework for promoting the sector, it has not yet been formally adopted. Consequently, India's policy landscape remains fragmented, with different aspects of data centre development addressed by policies on digital infrastructure, energy, taxation, and industrial development. Alongside MeitY's efforts, the TRAI has issued several recommendations to strengthen India's data centre ecosystem. The key policy initiatives are outlined below.

3.1 Existing Policies in India

3.1.1 Draft Data Centre Policy, 2020 (MeitY)

The **Draft Data Centre Policy** seeks to position India as a globally competitive data centre destination by establishing an enabling regulatory, infrastructural, and investment ecosystem. Recognising the strategic importance of data centres in supporting the digital economy, the policy adopts a comprehensive approach centred on five key pillars.

First, the policy seeks to enhance the **ease of doing business** by proposing measures such as granting infrastructure status to data centres, introducing a single-window clearance mechanism, providing fiscal incentives through a proposed Data Centre Incentivisation Scheme (DCIS), and facilitating the development of plug-and-play Data Centre Parks to streamline investments and reduce regulatory barriers.

Second, it focuses on creating an **efficient operational ecosystem** by ensuring access to reliable and affordable power, promoting the adoption of renewable energy, strengthening domestic and international connectivity, recognising data centres as essential services, and introducing dedicated building standards tailored to the unique operational requirements of data centre infrastructure.

Third, the policy envisages the development of **Data Centre Economic Zones (DCEZs)**, specialised clusters equipped with advanced digital, energy, and connectivity infrastructure. These zones are intended to attract hyperscale investments, foster economies of scale, and create integrated digital infrastructure ecosystems capable of supporting next-generation technologies.

Fourth, the policy emphasises **technology development, research, and capacity building** by promoting domestic manufacturing of Information Technology (IT) and non-IT equipment, encouraging research and innovation, supporting startups and MSMEs, and strengthening human capital through targeted skilling and workforce development initiatives.

Finally, the policy proposes a robust **institutional governance framework** comprising an Inter-Ministerial Empowered Committee (IMEC), a Data Centre Facilitation Unit (DCFU), and a Data Centre Industry Council (DCIC). Together, these institutional mechanisms are intended to coordinate policy implementation, facilitate stakeholder engagement, and ensure effective governance of India's evolving data centre ecosystem.

⁴⁷ There is a proposed National Data Centre Policy.

Collectively, these five pillars reflect the policy's objective of creating a conducive environment for investment, fostering technological innovation, strengthening digital infrastructure, and positioning India as a leading global hub for data centre development.

3.1.2 TRAI Recommendations on Data Centres (Consultation Paper)

The **TRAI** has further strengthened the policy discourse on data centres by publishing a consultation paper to enhance the long-term resilience, sustainability, and global competitiveness of India's data centre ecosystem. Building on the objectives of the Draft Data Centre Policy, TRAI's recommendations seek to address regulatory, infrastructural, and investment-related challenges that may impede the sector's growth.

A key recommendation is to **recognise data centres as strategic infrastructure** by granting them the status of *Vital Infrastructure*. Such a designation would acknowledge the critical role data centres play in supporting India's digital economy and essential services, while ensuring greater operational continuity, priority access to critical resources, and enhanced resilience against disruptions.

TRAI also advocates for **energy and infrastructure reforms** to support the sector's rapidly growing power requirements. It recommends recognising data centres as integrated energy assets capable of generating, storing, and managing captive renewable energy to improve energy security and accelerate the transition towards sustainable operations. In addition, TRAI proposes developing dedicated building codes and technical standards that reflect the specialised design, safety, and operational requirements of modern data centre infrastructure.

To improve the investment climate, TRAI recommends a range of **fiscal incentives and regulatory simplification measures**. These include the introduction of revenue-linked incentives and targeted tax benefits to enhance the financial viability of data centre investments, alongside the institutionalisation of a single-window clearance mechanism to streamline regulatory approvals, reduce compliance burdens, and facilitate faster project implementation.

Collectively, TRAI's recommendations complement the Draft Data Centre Policy by reinforcing the need for a stable regulatory framework, sustainable energy integration, and investment-friendly reforms to position India as a globally competitive hub for digital infrastructure.

3.2 International Policy Landscape

Several jurisdictions have moved beyond viewing data centres solely as commercial infrastructure and have begun regulating them as critical digital infrastructure through integrated frameworks spanning energy efficiency, cybersecurity, planning, sustainability, and data governance.

3.2.1 United Kingdom (UK)

The UK has adopted a strategic approach to governing data centres by recognising them as Critical National Infrastructure (CNI), reflecting their importance to national security and economic resilience. This recognition has been complemented by planning reforms that facilitate the development of strategically significant data centre projects, including streamlined approval processes and dedicated AI Growth Zones. Alongside infrastructure planning, the UK is strengthening cybersecurity obligations through amendments to its Network and Information Systems (NIS) regime, introducing mandatory risk management measures, incident reporting requirements, and regulatory oversight for large data centre operators.⁴⁸ Together, these measures seek to balance investment promotion with resilience, security, and operational continuity.

⁴⁸ Ostridge, H. (2026). *Cyber security and resilience (Network and Information Systems) bill* (Library Briefing Paper No. LLN-2026-0032). House of Lords Library. [https://researchbriefings.files.parliament.uk/documents/LLN-2026-0032/2026-0032-Cyber-Security-and-Resilience-\(Network-and-Information-Systems\)-Bill-LARGE.pdf](https://researchbriefings.files.parliament.uk/documents/LLN-2026-0032/2026-0032-Cyber-Security-and-Resilience-(Network-and-Information-Systems)-Bill-LARGE.pdf)

3.2.2 European Union

The European Union regulates data centres through an interconnected set of legislative instruments addressing sustainability, cybersecurity, AI governance, and data governance. The Energy Efficiency Directive (EED) requires larger data centres to publicly report performance indicators relating to energy consumption, water use, waste heat recovery, and operational efficiency, reflecting the EU's emphasis on sustainability and transparency. Cybersecurity obligations are governed through the NIS2 Directive, which classifies data centres as essential digital infrastructure and imposes mandatory cybersecurity and incident reporting requirements.⁴⁹ In parallel, the EU AI Act, Data Act, Data Governance Act, and General Data Protection Regulation (GDPR) collectively establish rules relating to AI deployment, cloud portability, trusted data sharing, and personal data protection, creating a comprehensive regulatory ecosystem that integrates infrastructure governance with broader digital policy objectives.

3.2.3 Singapore

Singapore illustrates a controlled, resource-aware model of data centre expansion. After temporarily pausing additional growth, it resumed capacity allocation through a competitive process and awarded approximately 80 MW in 2023 to proposals assessed against energy efficiency, decarbonisation, connectivity and wider economic contribution.⁵⁰ Its 2024 Green Data Centre Roadmap subsequently targeted at least 300 MW of additional capacity in the near term, with potentially another 200 MW enabled through green-energy deployments.⁵¹ In August 2026, Singapore provisionally allocated a further 200 MW to four operators. The selected projects committed to source more than half of their capacity from green energy, adopt liquid cooling and energy-efficient equipment, and locate their facilities within a low-carbon data centre park on Jurong Island.⁵²

Singapore's stance is therefore neither an unrestricted build-out nor a continuing moratorium. It treats data centre capacity as a scarce resource to be allocated against strategic, economic and sustainability performance. For India, the relevant lesson is to link approvals and incentives in resource-constrained clusters to measurable efficiency, grid readiness, green power and domestic economic spillovers.

⁴⁹ European Parliament, & Council of the European Union. (2022). *Directive (EU) 2022/2555 of the European Parliament and of the Council of 14 December 2022 on measures for a high common level of cybersecurity across the Union, amending Regulation (EU) No 910/2014 and Directive (EU) 2018/1972, and repealing Directive (EU) 2016/1148 (NIS 2 Directive)*. *Official Journal of the European Union*, L 333, 80–152. <http://data.europa.eu/eli/dir/2022/2555/oj>

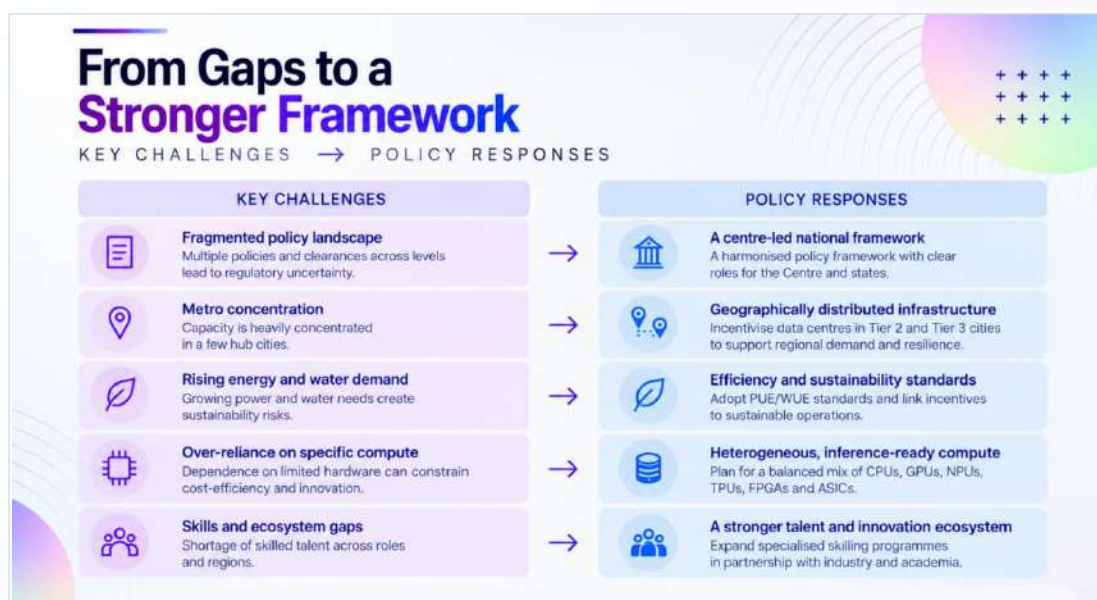
⁵⁰ Economic Development Board, & Infocomm Media Development Authority. (2023, July 14). *Four data centre proposals selected as part of pilot Data Centre-Call for Application*. <https://www.edb.gov.sg/en/about-edb/media-releases-publications/four-data-centre-proposals-selected-as-part-of-pilot-data-centre-call-for-application.html>

⁵¹ Infocomm Media Development Authority. (2024, May 30). *Singapore announces Green Data Centre Roadmap for sustainable growth of digital economy*. <https://www.imda.gov.sg/resources/press-releases-factsheets-and-speeches/press-releases/2024/sq-announces-green-data-centre-roadmap>

⁵² Economic Development Board, & Infocomm Media Development Authority. (2026, August 21). *Four data centre proposals selected as part of second Data Centre-Call for Application* [Fact sheet]. <https://www.edb.gov.sg/content/dam/edb-en/about-edb/media-releases/news/four-data-centre-proposals-selected-as-part-of-second-data-centre/Fact%20Sheet%20-Four%20data%20centre%20proposals%20selected%20as%20part%20of%20second%20data%20centre.pdf>

3.3 Key Policy Gaps in India

While India's Draft Data Centre Policy, 2020 and subsequent policy initiatives provide an important starting point, they remain primarily focused on investment promotion and ease of doing business. Compared to emerging international approaches, several important policy gaps remain.



3.3.1 Absence of a Comprehensive National Policy Framework

Although the Draft Data Centre Policy proposes several enabling measures, it has not yet been formally adopted, leaving India without a unified national framework for governing data centres. Consequently, the regulatory landscape remains fragmented across sectoral policies and state-level initiatives, resulting in varying regulatory approaches and limited coordination at the national level.

While state policies are necessary for attracting investment, the national framework should articulate a clear responsibility matrix. This allocation should reflect India's federal structure rather than attempt a watertight constitutional separation. Article 246 read with the Seventh Schedule of the Constitution places communications within the Union domain while land, local government and intra-state trade principally within the State domain. Electricity and acquisition of property constitute the Concurrent List.⁵³ The Centre should therefore set the national vision, common definitions, strategic compute priorities, baseline sustainability and cybersecurity standards, interoperability requirements, national reporting metrics, and arrangements for inter-state and international connectivity. States should lead project execution by identifying and facilitating land, coordinating local and municipal approvals, planning water and power availability, augmenting intra-state transmission and distribution capacity, structuring state incentives, and operating effective single-window systems. Since electricity and property acquisition are concurrent subjects, this allocation should operate on the basis of the modalities of a project and regional considerations. The Centre should provide direction and coordination, while states remain accountable for converting policy ambition into commissioned capacity. This approach would clarify and operationalise the institutional mechanisms already contemplated by MeitY and TRAI.^{54,55}

⁵³ India Const. art. 246, sched. 7, lists I-III.

<https://cdnbbsr.s3waas.gov.in/s380537a945c7aaa788ccfcdf1b99b5d8f/uploads/2024/05/20240516155637233.pdf>
<https://www.legislative.gov.in/static/uploads/2025/07/c9fe9c9b6840524844316f74bb1c556c.pdf>

⁵⁴ Ministry of Electronics and Information Technology. (2020). *Draft Data Centre Policy 2020*. Government of India. https://www.meity.gov.in/writereaddata/files/Draft%20Data%20Centre%20Policy%20-%2003112020_v5.5.pdf

⁵⁵ Telecom Regulatory Authority of India. (2022, November 18). *Recommendations on regulatory framework for promoting data economy through establishment of data centres, content delivery networks, and interconnect exchanges in India*. https://traigov.in/sites/default/files/2024-09/Recommendations_18112022.pdf

3.3.2 Limited Integration between Infrastructure and Energy Planning

Current policy proposals recognise reliable power supply as an important prerequisite for data centre development but stop short of providing a comprehensive framework for integrating data centres into long-term electricity planning. Issues relating to renewable energy integration, grid capacity, regional electricity demand, water availability, climate resilience, and sustainable resource management remain insufficiently addressed.

3.3.3 Absence of a Compute Strategy

Existing policy discussions primarily focus on attracting investment in physical data centre infrastructure, rather than positioning data centres as the foundational layer of India's broader AI and compute ecosystem. As a result, there is limited consideration of critical dimensions such as AI compute capacity, high-performance computing, heterogeneous compute architectures, cloud infrastructure, and long-term compute sovereignty.

Given evolving geopolitical dynamics, supply chain vulnerabilities, and constraints on access to advanced computing infrastructure, India needs to adopt a more strategic and forward-looking approach to computing. This should include promoting heterogeneous compute architectures tailored to India's specific requirements, such as local demand patterns, resource constraints, energy availability, and diverse AI workloads. Such an approach would enable India to build a more resilient, cost-effective, and sovereign compute ecosystem capable of supporting its long-term AI ambitions.

3.3.4 Limited Sustainability Standards and Reporting Frameworks

It is essential to move towards an ecosystem in which sustainability is embedded by design rather than treated as an afterthought. The government could prescribe standardised, and independently verifiable reporting requirements covering energy efficiency, carbon emissions, water consumption, waste heat utilisation, and lifecycle environmental impacts. These requirements should be accompanied by appropriately calibrated Power Usage Effectiveness (PUE) and Water Usage Effectiveness (WUE) ceilings, taking into account facility size, climatic conditions, and local water stress.

Fiscal incentives, long-tenor tax benefits, and relevant regulatory approvals should also be linked to measurable sustainability performance, including renewable-energy integration and the adoption of resource-efficient cooling and water-management solutions such as immersion cooling, recycled water, rainwater harvesting, and, where locally appropriate, desalinated seawater. Large facilities should additionally be subject to phased Energy Reuse Factor targets and waste-heat recovery obligations where technically feasible, enabling thermal by-products to be used in district energy systems or nearby industrial processes. This sustainability-by-design framework would provide stronger incentives for responsible investment, improve resource efficiency, and ensure that the growth of India's data-centre sector remains aligned with its long-term energy, water, and climate objectives.

3.3.5 Limited Cybersecurity Emphasis

Although cybersecurity is recognised as an important consideration, India's current policy framework could place greater emphasis on strengthening cybersecurity practices among data centre operators, drawing from approaches such as the EU's NIS2 Directive and the UK's evolving cybersecurity and resilience framework. Areas such as incident reporting, risk management, and sector-specific cybersecurity practices could be further strengthened to improve the overall resilience of the data centre ecosystem.

Therefore, rather than developing entirely new frameworks, it would be useful to engage ecosystem stakeholders to identify, adapt, and promote established cybersecurity best practices and international standards. This would enable India to strengthen the security and resilience of its data centre infrastructure while building on existing approaches and avoiding unnecessary duplication of frameworks.

3.3.6 Limited Focus on Access, Domestic Ecosystem Development, and Institutional Coordination

Current policies place considerable emphasis on attracting investment but devote comparatively less attention to ensuring affordable access to compute infrastructure for startups, MSMEs, researchers, and public institutions. Similarly, greater policy attention is required to strengthen domestic manufacturing, cooling technologies, semiconductor-linked supply chains, workforce development, and institutional coordination across digital infrastructure, energy, industrial policy, and AI governance.

4. Guiding Principles for a National Data Centre Policy

India's policy architecture must evolve to support the upcoming data centre build-out. A national framework is therefore overdue, but its value will lie less in the incentives it adds than in the coherence it imposes. Crucially, that framework must be oriented towards the workloads that will define the coming decade, ensuring the build-out is heterogeneous in its compute, inference-ready in its design, distributed across geographies, and capable of serving demand at the edge. These considerations are not confined to any single principle but thread through all of them. We have identified seven principles that should anchor that framework.

4.1 Strategic Autonomy

The first principle is that domestic compute must be treated as strategic infrastructure rather than as a purely commercial asset. India has committed roughly Rs. 10,372 crore to the IndiaAI Mission, whose shared compute pool has crossed 38,000 GPUs and is targeted to reach 100,000 by the end of 2026.⁵⁶ A policy built on this principle would treat the question of who owns, operates and can switch off critical compute as a matter of national resilience, hedging against supply-chain interference and foreign platform dependence without retreating into a closed economy. The balance to strike is between welcoming hyperscale and foreign direct investment, which India needs for scale and capital, and avoiding structural lock-in with few suppliers.

Strategic autonomy must also extend beyond the quantity of compute to its composition. A pool weighted towards training-optimised silicon leaves the country reliant on external capacity for the inference workloads that will dominate future demand, so the national interest lies in building heterogeneous, inference-ready capacity, spanning GPUs, CPUs, NPUs and purpose-built accelerators, able to serve real-time domestic demand without recourse to foreign platforms.

4.2 Heterogeneous, Inference-Ready Compute

The second principle is that compute capacity should be composed to match the workloads it is intended to serve, rather than defaulting to a single class of hardware. As established earlier, inference is expected to overtake both training and non-AI workloads and account for the largest share of data centre demand by 2030. However, inference-oriented capacity currently represents only a small share of India's installed AI compute, with the majority weighted towards training-optimised silicon procured for an earlier phase of the AI cycle. A national framework should therefore treat heterogeneity as a design objective in its own right, encouraging facilities that combine GPUs, NPUs, ASICs, and purpose-built inference accelerators. This would allow compute resources to be allocated according to the specific demand profile of the regions and workloads they serve. The benefits are mutually reinforcing: higher utilisation and efficiency, lower latency, and reduced power consumption, as inference-optimised hardware can deliver greater energy efficiency for suitable workloads than training-oriented hardware.

Historically, general-purpose CPUs formed the backbone of computing systems. The emergence of graphics processing units (GPUs) represented a significant shift towards large-scale parallel processing and substantially accelerated machine-learning workloads. More recently, specialised accelerators,

⁵⁶ Ministry of Electronics & IT. (2026, March 25). *IndiaAI mission expands AI ecosystem with affordable compute and startup support: Semicon India programme boosts domestic chip manufacturing and design ecosystem* [Press release]. Press Information Bureau, Government of India. <https://www.pib.gov.in/PressReleasePage.aspx?PRID=2245069&req=48&lang=2>

including NPUs, TPUs, field-programmable gate arrays (FPGAs), and ASICs, have emerged to optimise particular computational patterns and workloads.

This evolution reflects a fundamental principle: no single compute architecture is optimal for every workload. Training large foundation models requires high-throughput, massively parallel computing, which is typically enabled by GPUs and specialised accelerators. Inference workloads, particularly those deployed at scale or closer to the edge, may be more efficiently served by CPUs, NPUs, or other inference-optimised accelerators. Similarly, workloads across sectors such as telecommunications, finance, and industrial automation may require customised configurations that balance performance, cost, latency, and energy consumption. The literature increasingly indicates that improvements in compute efficiency depend not only on hardware specialisation but also on software optimisation, rather than simply scaling a single architecture uniformly. Industry analyses similarly suggest that heterogeneous systems can deliver gains in performance per watt, an increasingly important consideration as energy and resource constraints become more binding.

This shift is particularly relevant in the Indian context. India's AI ecosystem encompasses a diverse range of use cases, from large-scale foundation model development to AI deployments in resource-constrained environments such as rural healthcare, agriculture, and public service delivery. A homogeneous compute strategy risks over-provisioning infrastructure for some applications while under-serving others. Heterogeneous compute, by contrast, provides a pathway to align infrastructure more closely with the requirements of different workloads, sectors, and regions. The demand-allocation and empanelment mechanisms under the IndiaAI Mission therefore provide a ready policy lever and should be calibrated to support the growth of inference-ready, heterogeneous compute capacity, rather than defaulting predominantly to the procurement of training-oriented compute.

4.3 Sustainability by Design

The third principle is that environmental performance should be designed into facilities at the approval stage rather than retrofitted. Cooling is water-intensive in conventional data centre designs, and in a country attentive to water security this is a legitimate consideration. It is, however, increasingly a question that good design answers rather than aggravates. The sustainability of data centres should be assessed through a broader lens that considers energy efficiency, power sourcing, cooling design, and resource optimisation. Liquid and immersion cooling technologies sharply reduce both energy and water use, and they are moving from the experimental to the standard. The agreement in Madhya Pradesh to develop up to one GW of immersion-cooled, AI-ready capacity points to designs delivering substantial energy savings and water conservation of around ninety per cent relative to conventional approaches.⁵⁷ The Jamnagar facility's reliance on desalinated seawater rather than freshwater offers another template, decoupling cooling from stressed inland sources altogether.

Importantly, many efficiency improvements in the data centre sector are already driven by strong commercial incentives. Operators have a direct economic interest in reducing energy consumption, as inefficient cooling and power management significantly increase operational expenditure. PUE, therefore, is not merely an environmental metric but a key business performance indicator, with more efficient facilities achieving lower operating costs and greater competitiveness.

Furthermore, the rapid evolution of computing hardware naturally encourages continuous improvements in efficiency. AI and high-performance computing infrastructure undergo relatively short replacement cycles, often within two to four years, allowing operators to progressively adopt newer generations of processors, servers, and cooling systems that deliver higher computational output per unit of energy consumed.

The Union Budget 2026 has begun to operationalise this at the fiscal level, conditioning its long-tenor tax holiday for cloud and data centre operators, which runs to 2047, on compliance with sustainability and energy-efficiency standards covering power usage effectiveness, water conservation and responsible design.⁵⁸ This is a meaningful precedent as it establishes that fiscal incentives and environmental

⁵⁷ Submer. (2025, July 28). *Government of Madhya Pradesh signs strategic MoU with Submer for 1GW of sustainable, AI-ready datacenter capacity* [Press release]. <https://submer.com/blog/press/madhya-pradesh-signs-strategic-mou-for-ai-datacenters/>

⁵⁸ PIB Research. (2026, February 14). *Budget 2026–27 sets the stage for India as a global hub for cloud and AI infrastructure* [Backgrounder]. Press Information Bureau, Government of India. <https://www.pib.gov.in/PressReleasePage.aspx?PRID=2227953>

conditionality can be structured in tandem, and a national data centre policy should concretise this approach across the approval and licensing architecture rather than confining it to a single budgetary instrument. Equally important is the question of how data centres are classified under existing statutory frameworks. Data centres need to be classified as an essential service under the Energy Conservation Act, 2000, as well as under a national data centre policy. Several state-level instruments have already moved in this direction. Telangana's Data Centres Policy, for instance, designates data centres as an essential service, granting them twenty-four-hour operating privileges, exemptions from planned power cuts, and access to subsidised power tariffs. A national-level classification of this nature would provide regulatory certainty, reduce operational vulnerability to grid-level interruptions, and signal to investors that India treats digital infrastructure as critical infrastructure.⁵⁹ Singapore's competitive allocation of additional capacity provides a complementary example of how access to scarce infrastructure can be conditioned on sustainability, economic contribution and measurable resource efficiency.

The opportunity for India is to treat this moment not as a trade-off between digital ambition and environmental responsibility but as a chance to set a global benchmark for water-efficient, net-zero-aligned data centre design. This requires a policy approach that recognises where regulation can add value while allowing market mechanisms to continue driving innovation in areas such as cooling efficiency, hardware optimisation, and operational performance. At the same time, policymakers must focus on challenges that cannot be addressed solely through facility-level optimisation, including renewable energy integration, grid resilience, water availability, and long-term power planning. In this respect, two resource-level gaps in India's current framework demand urgent attention.

The first is water. Beyond energy, there are presently no national norms governing water consumption in data centre cooling systems. A national framework should establish water consumption intensity standards for new facilities, linked to regional hydrological assessments, and require operators to disclose water usage effectiveness alongside PUE as a standard performance metric. In water scarce regions, the increased demand can be met through policy-backed solutions such as rainwater harvesting and usage of recycled water from waste treatment plants. The government should also develop a standard compute utilisation matrix for data centres. Widely used indicators such as Power Usage Effectiveness primarily measure the efficiency of supporting infrastructure, including cooling and power distribution, but do not establish whether the underlying CPUs, GPUs, NPUs and other computing resources are being productively utilised. A technology-neutral framework could report provisioned capacity, average and peak utilisation, idle capacity, energy consumed per unit of useful computation and utilisation by compute category. This would help identify over-provisioned or stranded capacity, enable comparison across facilities and support more informed infrastructure planning.

The second gap concerns waste heat. Data centres produce significant quantities of produce significant quantities of waste heat that are currently not recovered in India, representing both an environmental cost and a missed resource opportunity. Internationally, this is already being addressed through regulation and commercial partnership. In Finland, Microsoft and energy utility Fortum have structured an arrangement whereby waste heat from data centre operations supplies approximately forty per cent of district heating demand for around 250,000 residents.⁶⁰ Germany has gone further, legally requiring new data centres to achieve a minimum Energy Reuse Factor that rises progressively to twenty per cent by 2027. India has no equivalent obligation, nor even a voluntary target.⁶¹ A national data centre policy could introduce a phased Energy Reuse Factor requirement for new large-scale facilities, accompanied by planning guidance encouraging co-location with district energy systems.

4.4 Distributed Infrastructure

The fourth principle is systemic geographic distribution of infrastructure. Supply today is heavily clustered, with the Mumbai and Chennai corridors accounting for close to seventy per cent of capacity. This

⁵⁹ Government of Telangana, Information Technology, Electronics & Communications Department. (2016). *Data centres policy*. <https://invest.telangana.gov.in/wp-content/uploads/2024/07/Telangana-Data-Centres-Policy-1.pdf>

⁶⁰ aquatherm GmbH. (2025, August 27). *Using waste heat from data centres: Turning digital heat into community warmth*. <https://blog.aquatherm.de/en/using-waste-heat-from-data-centres-turning-digital-heat-into-community-warmth>

⁶¹ Haroldsen, J. O. (2026, May 25). German stakeholders share insights on using data center waste heat in district heating at European Tech Show. *Natural Refrigerants*. <https://naturalrefrigerants.com/news/german-stakeholders-data-center-waste-heat-district-heating/>

concentration is driven primarily by economic logic. Hyperscalers locate cloud regions where enterprise demand, population density, network connectivity, and latency requirements converge, reinforcing the gravitational pull of cities with established digital ecosystems. This pattern concentrates grid loading and connectivity risk in two regions, leaving the national network exposed to localised shocks. A national framework should actively reward dispersal. This includes tiered incentives for tier 2 and tier 3 locations, support for edge facilities serving low-latency, real-time AI applications, and coordination with transmission planning so that capacity follows available power rather than straining a few overloaded corridors.

Distribution also reflects the spatial implications of the shift towards inference. Because inference is latency-sensitive and best served close to the user, dispersal allows compute to be matched to regional demand, with the heterogeneous mix of a given facility tuned to the workloads of the area it serves. For example, NPU-heavy capacity can be located near clusters of healthcare or public-service demand. The migration of edge capacity to cities such as Jaipur, Visakhapatnam, Kochi, Lucknow, etc., should be consolidated by policy rather than left to market action. It must, however, be acknowledged that tier 2 and tier 3 cities do not yet generate sufficient enterprise demand to justify large GPU clusters. The policy imperative is therefore such: large-scale, GPU-heavy hyperscale facilities should be directed towards regions with demonstrated enterprise demand and infrastructure readiness, while purpose-specific, distributed edge facilities optimised for inference workloads are deployed at regional scale to build local data ecosystems and extend market reach. Distribution serves three ends at once, namely resilience, more balanced regional development, and relief for the grids of the established hubs.

4.5 Interoperability and Open Standards

The fifth principle responds to a risk that efficiency alone cannot resolve. Much of the prevailing inference toolchain is tightly coupled to a narrow set of hardware and stacks, so that optimisation achieved on one vendor's architecture does not port cleanly to alternatives, and gains in utilisation can deepen rather than loosen external dependence. Left unaddressed, this is the point at which efficiency and sovereignty pull apart. A national framework should therefore favour open standards, portable orchestration layers and interoperable tooling that let inference workloads move across heterogeneous hardware, including domestic accelerators, without prohibitive re-engineering. Interoperability is the connective principle that allows the others to hold, since heterogeneous compute, a distributed edge estate and strategic autonomy are each undermined if workloads remain captive to a single supplier's stack.

4.6 Access, Affordability, and Ecosystem Development

Access, Affordability, and Ecosystem Development as a principle of national data centre strategy rests on the foundational recognition that compute infrastructure, if concentrated only among large enterprises, becomes a tool of exclusion rather than growth, and that democratising access to it is as important a policy objective as building the infrastructure itself. India has moved in this direction through the IndiaAI Mission, approved in March 2024 with an outlay of Rs. 10,371.92 crore over five years, making over 38,000 high-end processors available to startups, researchers, and public institutions at Rs. 65 per hour, roughly one-third of the global average. Public initiatives such as GI Cloud MeghRaj and the National Supercomputing Mission broaden access to cloud services and high-performance computing resources, enabling government departments, academic institutions and researchers to use computing capacity that they may be unable to develop independently. The AIKosh platform adds to this by aggregating over 7,500 datasets and 273 AI models as shared national resources, so startups can develop AI applications without collecting primary data from scratch.⁶²

On the supply side, India remains heavily import-dependent for equipment and advanced cooling technologies, which are increasingly critical as AI workloads demand liquid-to-chip or immersion-based cooling systems. A dedicated Production-Linked Incentive (PLI) scheme for data centre equipment and specialised manufacturing clusters under Make in India would help close this gap and reduce structural vulnerability to global supply chain disruptions.

⁶² Press Information Bureau. (2025, October 12). *Transforming India with AI* (Release ID: 2178092). Ministry of Information and Broadcasting, Government of India.
<https://www.pib.gov.in/PressReleasePage.aspx?PRID=2178092®=48&lang=2>

The global demand for skilled data centre workers presents an additional opportunity. Meta has committed 115 million United States dollars to train workers in electrical work, mechanical systems, plumbing, and fibre installation, covering all costs and guaranteeing jobs upon completion, reflecting a shortage of hundreds of thousands of skilled tradespeople globally. India is already demonstrating it can meet this demand, although there remains room for improvement.⁶³

4.7 Skill & Workforce Development

Skill & Workforce Development has become one of the most important principles of India's national data centre strategy because the country is building infrastructure faster than it is building the people needed to run it. According to the Economic Survey 2025-26,⁶⁴ India's data centre capacity is on a trajectory to reach 8 GW by 2030, a massive jump from the 1.4 GW baseline recorded as of mid-2025, and yet the sector currently employs only around 90,000 professionals, with specialists in roles such as facilities engineering, IT operations, network engineering, cloud operations, cybersecurity and infrastructure management remaining in severe shortage. While individual data centres create relatively few operational jobs, their broader economic impact is considerably larger. Data centres stimulate specialist engineering expertise, server management, infrastructure maintenance, networking services, cloud ecosystem development, IT services, managed service providers and export-oriented technical services. Moreover, cloud regions often coincide with growing technology ecosystems and Global Capability Centres (GCCs), reinforcing regional digital economies. Future workforce planning must account for this opportunity and address the demand-supply gap currently observed in the AI and data ecosystem.

The talent shortage is particularly acute in AI infrastructure operations. The report assigns this segment a Supply Sufficiency Index score of 47, indicating that the availability of qualified professionals falls considerably short of the workforce required to manage increasingly AI-intensive computing environments.⁶⁵ The problem is not simply one of numbers but of quality and specialisation. The MSDE's National Skill Gap Study 2025⁶⁶ identifies that India's existing training systems suffer from poor alignment between what institutions teach and what data centre employers actually need, with in-demand job roles now including Data Engineering, Cloud Computing, Cyber Cloud Security, Natural Language Processing, and LLM development, none of which are adequately covered in conventional engineering or IT curricula. Making this more complex is the physical nature of data centre work itself, because as AI workloads intensify, data centres are shifting from conventional air-cooled systems designed for 12 to 18 kilowatt workloads to direct liquid-to-chip cooling systems capable of handling up to 150 kilowatts and above, meaning that cooling and power technicians must now be as tech-savvy in software and operations as they are in mechanical systems, a combination that is extremely rare in the current workforce.

India's skilling strategy must transition from broad digital literacy to highly specialised, industry-led technical training. The government should collaborate directly with hyperscalers, telecommunication providers, and semiconductor firms to build academic curricula tailored to data centre management. The workforce development pipeline must address both ends of the digital spectrum. It should generate high-end engineering roles in hardware design and infrastructure maintenance, alongside mass employment opportunities in data annotation, geospatial analysis, medical diagnostics, and localised AI deployment services.

The government has responded through the Skill India Mission, under which Pradhan Mantri Kaushal Vikas Yojana (PMKVY) has trained 7,25,517 candidates, National Apprenticeship Promotion Scheme (NAPS) has engaged 4,35,915 apprentices, and Craftsmen Training Scheme (CTS) through ITIs has trained 7,55,736 workers in the Information Technology and Information Technology-enabled Services (IT-ITeS) sector, including data centres, cybersecurity, and AI/ML, while platforms like the Skill India Digital Hub and MeitY's FutureSkills Prime provide online access to industry-relevant courses. At the specialised end, NITI Aayog

⁶³ Munis, J. (2026, June 10). A 5-week course and a guaranteed job: Meta commits \$115 million to solve the skilled-trades shortage stalling its AI buildout. *Fortune*. <https://fortune.com/2026/06/10/meta-data-center-free-training-employment-blue-collar-trade-jobs/>

⁶⁴ Department of Economic Affairs. (2026, January). Economic survey 2025-26. Ministry of Finance, Government of India. <https://www.indiabudget.gov.in/economicsurvey/doc/echapter.pdf>

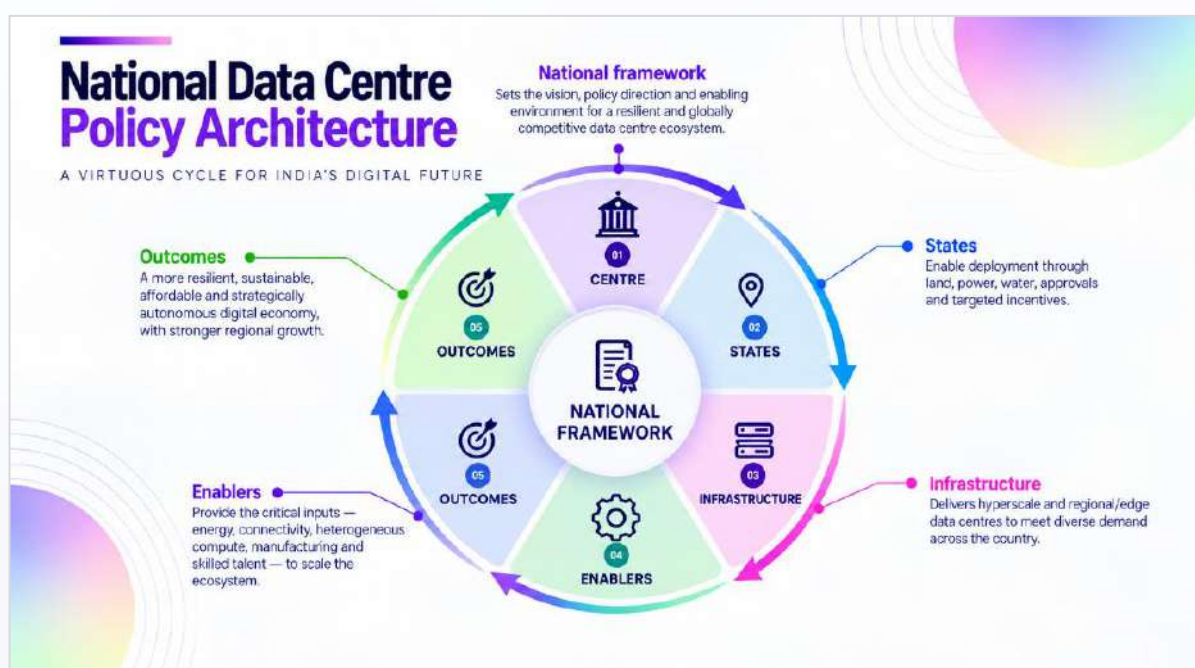
⁶⁵ Bagchi, S. (2026, March 11). India's \$13-billion data centre boom faces AI talent crunch. *TechCircle*. <https://www.techcircle.in/2026/03/11/india-s-13-billion-data-centre-boom-faces-ai-talent-crunch/>

⁶⁶ National Council of Applied Economic Research. (2025). *National skill gap study for high growth sectors*. Ministry of Skill Development and Entrepreneurship, Government of India. <https://www.msde.gov.in/static/uploads/2025/07/6a62ad4129b524c392ed1450393804f4.pdf>

(National Institution for Transforming India)'s Technology Services report of February 2026 recommends building a Unified AI Skilling Engine integrated with NAPS and PMKVY that delivers role-based upskilling and micro-credentials, and proposes an India AI Talent Mission to address the widening shortage in next-generation profiles like GenAI researchers and Large Language Model Operations (LLMOps) specialists.⁶⁷ On the industry-academia side, STT GDC India in collaboration with Schneider Electric India Foundation set up a dedicated skilling centre in Bengaluru specifically to address the entry-level workforce shortage in the data centre sector, with a curriculum developed by Singapore Polytechnic under an MoU with National Skill Development Corporation (NSDC), representing the kind of direct, practical, equipment-based training that the industry needs far more of.⁶⁸

A particularly important geographic dimension is that data centres are no longer concentrated only in Mumbai, Delhi-NCR, and Chennai but are decentralising into Tier 2 and Tier 3 cities, and as NITI Aayog notes, cities like Jaipur and Kochi are already emerging as GCC hubs, which means the workforce pipeline must be built in those regions too, through regional ITIs, polytechnics, and locally embedded training partnerships, because a data centre in a smaller city cannot depend on constant supervision or talent transfer from a central hub and needs independently capable, locally trained operators.⁶⁹

Importantly, the success of these skilling initiatives should be assessed through outcome-based metrics rather than programme-delivery metrics. The number of candidates enrolled or courses completed should not, by itself, be treated as evidence of success. Instead, evaluation should focus on indicators such as job placement, employment retention, wage progression, and the extent to which training translates into industry-relevant employment. In other words, the key metric should be how many individuals secure and sustain relevant employment as a result of these programmes, rather than simply how many individuals participate in them.



⁶⁷ Press Information Bureau. (2026, January 23). *Year end review 2025: Ministry of Skill Development and Entrepreneurship*. Ministry of Skill Development and Entrepreneurship, Government of India.

⁶⁸ STT GDC India. (2021, July 27). *Bridging the skill gap in the data centre industry*.

⁶⁹ NITI Frontier Tech Hub. (2026, February). *India's technology services: Reimagination ahead*. NITI Aayog, Government of India.

<https://niti.gov.in/sites/default/files/2026-02/Technology-Services-Reimagination-Ahead.pdf>

5. Key Recommendations and Implementation Roadmap

Moving from a fragmented, state-level approach to a cohesive national strategy requires a coordinated framework. The recommendations below align with the guiding principles of strategic autonomy, heterogeneity, sustainability, and distributed architecture, and are followed by a phased implementation roadmap.

5.1 Governance and Regulatory Alignment

- **Notify a Centre-led National Policy:** MeitY should define the national vision, common standards and outcome targets for compute, sustainability, cybersecurity, connectivity and resilience. The framework should establish common definitions and reporting metrics while preserving states' discretion over locally appropriate execution and incentives.
- **Establish a Federated Single-Window System:** Connect the national portal to empowered state nodal agencies rather than attempting to centralise local approvals. States should remain accountable for land and building approvals, utility connections, grid augmentation, water planning and other time-bound operational clearances, while the Centre addresses inter-ministerial and inter-state bottlenecks.

5.2 Infrastructure and Spatial Planning

- **Incentivise Tier-2 and Tier-3 Distribution:** Provide tiered financial incentives to encourage the spatial distribution of purpose-built edge facilities into emerging hubs (e.g., Jaipur, Visakhapatnam, Kochi, Lucknow). This can help reduce grid strain arising from the concentration of data centre capacity in the Mumbai and Chennai corridors and support regional data ecosystems.
- **Align with Localised Workloads:** Locate edge compute nodes to areas with high demand for latency-sensitive applications, such as regional healthcare institutions and public service networks, to process data closer to its source, reduce inference latency, and comply with applicable requirements under the DPDP Act and sector-specific data-residency rules.
- **Embed Cyber Resilience in Spatial Design:** Ensure that distributed edge deployments are accompanied by location-specific security architectures to protect the expanded attack surface.

5.3 Compute Architecture and Access

- **Incentivise Heterogeneous Compute:** Direct procurement frameworks and incentives (such as the IndiaAI Mission) to support inference-ready, heterogeneous architectures (blending GPUs, NPUs, ASICs, TPUs, FPGAs, and AI-accelerated CPUs) rather than defaulting exclusively to GPUs.
- **Mandate Interoperability and Open Standards:** Enforce open-standard software frameworks and hardware interoperability across ecosystems to ensure workloads can move seamlessly without restrictive re-engineering.

5.4 Sustainability and Energy Integration

- **Standardise PUE and WUE Metrics:** Integrate PUE and WUE ceilings.
- **Condition Fiscal Incentives on Sustainability:** Mandate renewable energy integration and resource optimisation (such as utilising desalinated seawater or immersion cooling) as a condition for long-tenor tax holidays and licensing.
- **Introduce Waste Heat Recovery Obligations:** Develop planning guidance and phased Energy Reuse Factor targets for large-scale facilities to enable waste heat recovery, allowing thermal by-products to be reused in district energy systems or for local industrial purposes.

5.5 Workforce and Ecosystem Development

- **Incentivise Domestic Manufacturing (PLI Scheme):** Introduce a dedicated PLI scheme for data centre equipment, semiconductor-linked supply chains, and advanced cooling technologies to reduce reliance on imported hardware.
- **Launch a Unified AI Skilling Engine:** Implement NITI Aayog's recommended India AI Talent Mission and Unified AI Skilling Engine to bridge the talent gap in critical fields such as cloud security, LLMOps, data engineering, and liquid-cooling HVAC maintenance.
- **Localise the Talent Pipeline:** Expand specialised data centre training pipelines into Tier-2 and Tier-3 cities through regional Industrial Training Institutes (ITIs), polytechnics, and industry-academia partnerships to ensure decentralised facilities are supported by an independently capable, local workforce.

5.6 Implementation Roadmap

The transition to this strategic policy framework should be executed through a phased rollout to balance immediate investment facilitation with long-term sovereign and sustainable scaling.

Phase	Timeframe	Strategic Objectives and Actions
Phase 1: Immediate Actions	0 to 12 Months	1. Notify a Centre-led National Data Centre Policy defining common standards, targets and reporting metrics. 2. Establish a central nodal agency and empowered state facilitation units with a published responsibility matrix. 3. Connect national and state single-window portals through time-bound, trackable approval processes. 4. Direct the IndiaAI Mission towards empanelling heterogeneous compute architectures.

Phase 2: Structural Rollout	1 to 3 Years	1. Roll out tiered financial incentives specifically for Tier 2 and Tier 3 distributed edge locations. 2. Incentivising sovereign compute development by leveraging National Supercomputing Mission (NSM) 2. 3. Integrate mandatory PUE and WUE ceilings into national building codes. 4. Deploy a Unified AI Skilling Engine across regional technical institutes (ITIs/polytechnics) to localise talent pipelines. 5. Introduce sector-specific cybersecurity and incident reporting standards.
Phase 3: Sovereign Scaling	3 to 5 Years	1. Enforce national interoperability and open standards. 2. Mandate minimum renewable energy procurement and phased heat-reuse targets (Energy Reuse Factor) for all new facilities. 3. Achieve national targets for subsidised, heterogeneous compute access for startups and MSMEs via the IndiaAI Mission. 4. Mature domestic supply chains through operational PLI schemes for advanced data centre cooling and networking equipment.

6. Conclusion

India stands at a critical juncture in the development of its data centre ecosystem. The rapid expansion of AI, cloud computing, DPI, and data-intensive services will make data centres increasingly central to India's economic competitiveness and technological ambitions. Yet the current policy landscape remains fragmented and predominantly oriented towards investment facilitation and physical infrastructure expansion. As this paper demonstrates, the next phase of growth requires a shift from viewing data centres simply as commercial infrastructure to recognising them as strategic infrastructure that underpins India's AI, compute, energy, and digital economy ambitions.

A National Data Centre Policy should therefore provide the coherence that the current ecosystem lacks, while preserving states' ability to compete for investment and tailor incentives to their local conditions. The objective should not be to centralise data centre development, but to establish a common national direction covering compute, energy, sustainability, cybersecurity, connectivity, workforce development, and ecosystem growth. In particular, India's approach to compute must anticipate the transition towards inference-intensive workloads by encouraging heterogeneous, inference-ready architectures, open standards, and interoperable infrastructure rather than allowing the ecosystem to become dependent on a narrow set of hardware and technology stacks. Under this model, the Centre would provide strategic direction and policy coherence, while states would translate that direction into commissioned capacity through land facilitation, power and water planning, local approvals, investment promotion and ecosystem development.

Equally, the next phase of expansion must be sustainable by design and geographically distributed. Data centre growth will place increasing demands on electricity, water, and other critical resources, making coordinated energy and infrastructure planning essential. At the same time, a combination of hyperscale facilities in infrastructure-ready locations and distributed, inference-oriented edge capacity in Tier 2 and Tier 3 cities can improve resilience, reduce latency, distribute supply, and support more balanced regional development. Sustainability, compute architecture, and spatial planning should therefore be treated not as separate policy questions, but as interconnected components of a national infrastructure strategy.

The policy framework must also ensure that the benefits of this infrastructure extend beyond large hyperscalers and enterprises. Affordable access to compute, domestic ecosystem development, interoperable technologies, advanced cooling technologies and domestic equipment manufacturing, and industry-aligned workforce development will be critical to enabling startups, MSMEs, researchers, and public institutions to participate in the emerging compute economy. The proposed roadmap consequently envisages a phased transition, from establishing a unified national framework and streamlined approvals in the immediate term, to embedding sustainability, regional skilling, cybersecurity, interoperability, and domestic supply chains over the medium term.

Ultimately, India's objective should not simply be to build more data centres, but to build the right data centres, in the right places, powered by the right energy systems, using the right compute architectures, and connected to the wider domestic technology ecosystem. A forward-looking National Data Centre Policy can provide the foundation for this transition, enabling India to convert the coming surge in data and AI demand into a source of economic growth, technological resilience, strategic autonomy, and sustainable digital development.

Authors



Soham Jagtap

Senior Research Associate, The Dialogue

Soham Jagtap is a technology policy professional with an LLM in Law and Technology and over two years of experience at The Dialogue. Works at the intersection of AI governance, fintech regulation, digital public infrastructure, and data protection, producing research and advisory outputs that engage senior stakeholders across the RBI, government, and international institutions. Combines rigorous analytical writing with direct experience convening high-level policy dialogues, synthesising complex regulatory terrain into decision-ready outputs.



Anuradha Bhattacharya

Senior Research Associate, The Dialogue

Anuradha Bhattacharya is a Senior Research Associate at The Dialogue, where she works at the intersection of competition law, digital markets and emerging technologies. Her work spans AI and competition, AI sustainability, data-centre policy, responsible AI, and the implications of emerging technologies for the future of work and platform economies. She has led research and policy initiatives involving stakeholder consultations, legislative and regulatory engagement, and support to Members of Parliament, including work examining how AI and technology are reshaping markets, businesses, employment and digital ecosystems. She also led the development of the RAISE Index on responsible AI in collaboration with the Government of Telangana, which was launched at the World Economic Forum in Davos. Prior to joining The Dialogue, she was a Project Fellow at the Vidhi Centre for Legal Policy, where she contributed to the Ministry of Corporate Affairs' Committee on Digital Competition Law and the drafting of the Report of the Committee on Digital Competition Law released in 2024, as well as research on corporate law and financial regulation. She holds an LL.M. in Competition Law from University College London (UCL), where she graduated with Distinction, and a B.A. LL.B. (Hons.) from the University of Calcutta.



Kamesh Shekar

Associate Director - Strategy & Research, The Dialogue

Kamesh Shekar is Associate Director – Strategy & Research at The Dialogue, where he leads the organisation's corporate affairs and external engagement strategy and drives research and policy initiatives across emerging technology domains, including artificial intelligence, semiconductors, hardware manufacturing, data centres, privacy, and data governance. He also serves as a member of Elsevier's Responsible AI Advisory Board, contributing to thought leadership and practical frameworks for the safe and effective adoption of AI. At The Dialogue, Kamesh directs the Coalition for Responsible Evolution of AI (CoRE-AI), a multi-stakeholder coalition of over 55 members spanning the AI value chain, focused on advancing responsible AI development and policy dialogue.



thedialogue.org.in



LinkedIn | The Dialogue



X | The Dialogue



Whatsapp | The Dialogue



Instagram | The Dialogue