PAHLE INDIA FOUNDATION

{X}KDR

AI KNOWLEDGE CONSORTIUM

Assessing ‘Compute’ Capacities and Policy Frameworks for Artificial Intelligence in India

MAY 2025

Assessing ‘Compute’ Capacities and Policy Frameworks for Artificial Intelligence in India

MAY 2025

PAHLE INDIA FOUNDATION

Authors: Rajiv Kumar, Ajay Shah, Abhishek Jha, Ankeetaa Mahesshwari, Garima Jain, Karthik Suresh

In this paper, we assess the status of ‘compute’ capacities for artificial intelligence (AI) in India. We take stock of recent policy developments relating to ‘compute’ in India as well as in other countries. We then discuss a few policy options that would incentivize the production and availability of these resources in India. We emphasize on the need for reforms in procurement, as well as the need for foreign policy measures that ensure the uninterrupted supply of the resources required for ‘compute’. Our observations, while specifically focusing on the element of ‘compute’ in the AI triad, nevertheless have broader implications when thinking about policy options for all kinds of inputs needed to ensure India’s continued importance in the global market for information technology services.

JEL Codes: *L8, K2, O3*

Dr. Rajiv Kumar, Garima Jain, Abhishek Jha, and Ankeetaa Mahesshwari are researchers at Pahle India Foundation, New Delhi, India. Ajay Shah and Karthik Suresh are researchers at XKDR Forum, Mumbai, India. All authors contributed equally. We thank Tushar Anand, Sudipto Banerjee, Vikash Gautam, Venkatesh Hariharan, Harleen Kaur, Pranay Kotasthane, Ayush Patnaik, Ravi Pokharna, Siddarth Raman, Vivan Sharan, Susan Thomas, and Niharika Yadav for useful discussions. All errors and omissions remain our own.

Introduction

The term "artificial intelligence" (AI) denotes a set of general-purpose technologies that have the potential to deeply transform economies and societies (Brynjolfsson et al., 2019). The Organisation for Economic Co-operation and Development (2019) (OECD) defines AI as "machine-based systems that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions influencing real or virtual environments". Three kinds of inputs are needed to set up AI systems. These three inputs are collectively referred to as the "AI triad" (Buchanan, 2020; Sastry et al., 2024). The AI triad consists of:

1. **Data**: Machine learning systems learn about the world by using structured or unstructured data sets.[1]

2. **Algorithms**: An algorithm is a series of instructions on how the computer should process information. When it comes to AI, algorithms make inferences from data and feedback using various techniques. There are three broad approaches to machine learning : (i) supervised learning, (ii) unsupervised learning, and (iii) reinforcement learning (Buchanan, 2020).

3. **"Compute"**: AI relies on "physical hardware and software infrastructure supporting AI workloads, including one or more 'stacks' (layers) of hardware and software used to support specialised AI workloads and applications in an efficient manner". This is referred to as "AI Compute", "compute capacity", or simply "compute" (Organisation for Economic Co-operation and Development, 2023). "Compute" in turn rests on three kinds of inputs: (i) hardware for "compute" include the appropriate chips, and graphics processing units (GPUs); (ii) infrastructure includes data centres, servers, etc.; and (iii) software includes general purpose computing for GPUs like OpenCL, CUDA, and the frameworks that depend on them e.g. PyTorch, TensorFlow, etc.

AI systems require all three elements of the triad to work well together. The following statement explains their inter-dependence: "Machine learning systems use computing power to execute algorithms that learn from data" (Buchanan, 2020).

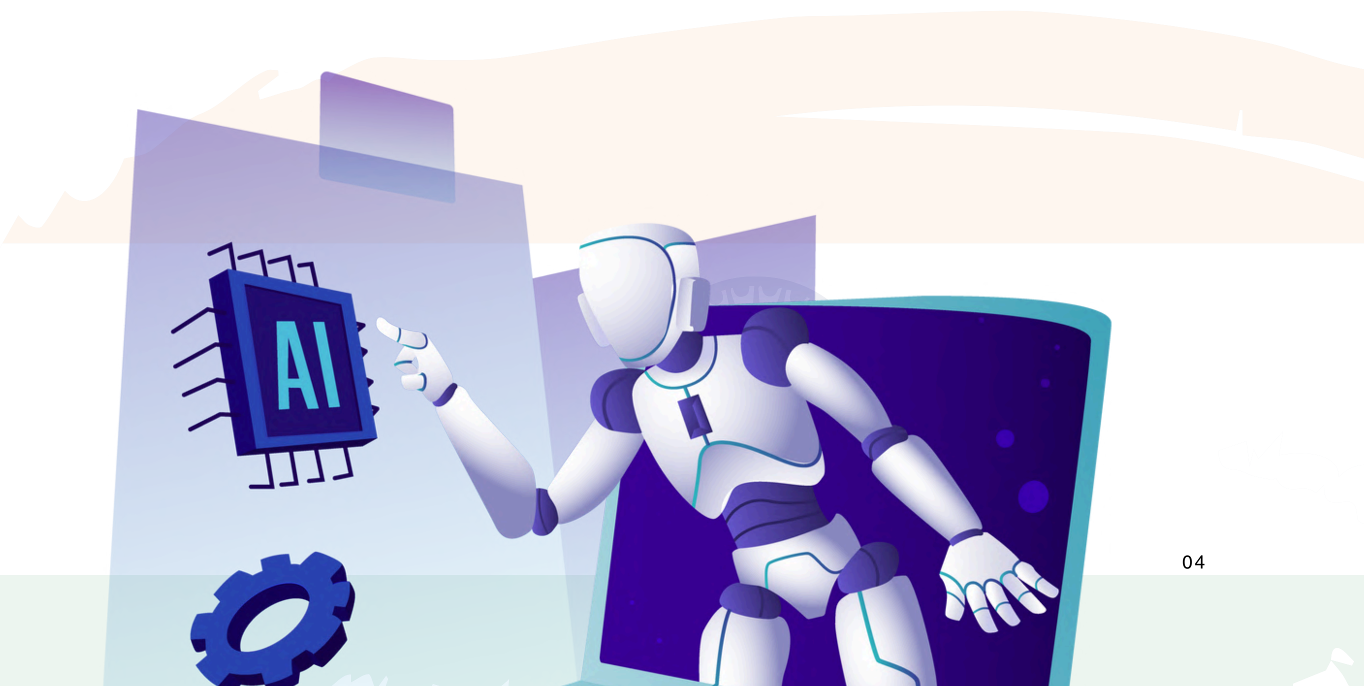
"Compute" for AI, in terms of policy thinking and options, has been the least understood element of the triad (Organisation for Economic Co-operation and Development, 2023; Vipra & West, 2023). "Compute" for AI requires resources that facilitate high-performance computing (HPC). However these resources are increasingly being "securitised", i.e., major countries are treating their "compute" technology with the seriousness of a strategic resource (Kotasthane & Manchi, 2023). As a consequence, the supply of hardware and even infrastructure for "compute" has been restricted by some countries which are also the largest producers and consumers of these inputs. In this new reality, we as Indian researchers present our attempt to understand the resources that are required to establish the computational needs for AI. We then discuss a few policy options that would incentivize the production and availability of these resources in India. Our observations, while specifically focusing on the element of "compute" in the AI triad, nevertheless have broader implications when thinking about policy options for all three elements of the AI triad.

Why is "compute" important in the Indian context?

The IT sector is India's largest and most productive industrial sector (see Appendix 1).(For further reading, see Shah, 2022) India's success in the IT sector can be attributed primarily to the comparative advantages from a low-cost, highly-skilled English-speaking workforce (Singh, 2016), buoyed by the positive externalities generated by India's strong higher and tertiary education systems, etc. We also note the role played by government policies which reduced capital and operational costs for IT firms. These policies have included fiscal incentives to the IT sector by both Union and state governments, e.g., amendments to sections 10A and 10B of the Income Tax Act in the year 2000; and industrial policies, e.g., Karnataka's Electronic System Design and Manufacturing (ESDM) policy (2013) and New Investment Incentive Policy (2013). These factors incentivised private sector investment, particularly into firms that provide low-cost software development and IT services (Das & Sagara, 2017).

Around the year 2010, *deep learning* emerged as an important paradigm in machine learning. The accuracy of a deep learning model is directly proportional to scale. This requires vast levels of computational power i.e., "compute". Accordingly, the Indian IT sector identifies AI as a critical area of growth, especially when it comes to "human capital" and "compute" (NASSCOM, 2024a). Companies are building their capabilities to service the AI applications space, as well as building tools for automation, data, and analytics solutions (Jujjavarapu et al., 2018; NASSCOM, 2024a). One particular area of interest is the role of global capability centres (GCC) when it comes to R&D in AI (NASSCOM, 2024b).

When it comes to the human capital that develops algorithms and manages data, India seems be well-positioned to service the AI revolution. Nearly 30% of the global AI talent pool (machine learning engineers, data scientists, AI specialists etc.) hold, or at some point have held, Indian citizenship (Hemmen & Pal, 2024). In fact, a group of researchers from the Indian Institute of Science in Bengaluru, along with their counterparts from the University of Michigan in the United States, won the prestigious Gordon Bell prize in 2023 for their research on quantum accuracy using the world's fastest supercomputer, the *Frontier* (Indian Institute of Science, 2023). No doubt, these achievements signify the impressive progress that Indians could achieve with high-performance computing. However, the shortfall lies with "compute" and the "invisible infrastructure" that surrounds it.[2]



What is the current state of “compute” in India?

In this section, we discuss the existing set of computers that are capable of delivering high-performance operations in India.

Supercomputers are high-performance computers that can deliver highly intensive computing solutions quickly and at scale. The terms “supercomputing” and “high performance computing” are used synonymously in technical literature (Sterling et al., 2017) as well as legal texts such as export regulations. For example, under United States law, a supercomputer is a “computing system having a collective maximum theoretical compute capacity of 100 or more double-precision (64-bit) petaflops or 200 or more single-precision (32-bit) petaflops within a 41,600 cubic feet or smaller envelope”.[3]

Historically, in India, supercomputers were developed for military and scientific use. Initially, these supercomputers had to be developed indigenously since the United States (US) and other NATO member states had restricted the sale of high-performance computers to India (Subramanian, 2021; Wolcott et al., 1998). Therefore, in 1991, the Center for Development of Advanced Computing (C-DAC) released the *Param-8000* supercomputer. It is interesting to note that, as a consequence of the limited available parts that were not restricted in the import markets, Indian supercomputers used parallel computing — a technology that today’s fastest GPUs also use (Mashelkar, 2008; Sanglard, 2020). Later versions of the *Param* model were improved upon and reformed with greater and greater processing speeds. The Param *Padma*, introduced in 2003, was the first Indian computer to break into the HPC Top500, a biannual publication that compiles the list of the fastest computers in the world. Subsequent iterations such as the Param *Yuva* and the Param *Siddhi* have also been featured in the Top 500. Tata Consultancy Services Ltd. also developed a supercomputer known as “Eka” which was released in 2008 and held a spot in the Top500 for some time. Eka was built entirely with funding from the Tata group of companies. All of India’s supercomputers developed so far — including the “Eka” — are used for R&D purposes with only a few exceptions. These uses range from weather detection and monitoring, to the determination and construction of molecules in chemistry and biochemistry research, etc.

Until recently, Indian supercomputers had not been used in R&D for AI. This has changed with the introduction of Param *Siddhi-AI* (2020) and Param *Airawat* (2023). In the latest iteration of the Top 500 list, there are six Indian computers — *Airawat, Arka, Arunika, Arka-AI, Pratyush* and *Mihir*. All of them are housed in, and operated by, research institutes of the Government of India.

In the previous decade, “deep learning” gained popularity because it was discovered that a cluster of specialized processors include graphical processing units (GPUs), tensor processing units (TPUs), neural processing units (NPUs) etc. This enabled the transition away from general-purpose CPUs to specialized processors which more computation per unit time with less energy. Until then, GPUs were known in the computer industry for their ability to provide high-quality outputs to enhance gaming and multimedia consumption. A “cluster” of GPUs can now provide high levels of computing capacity. Indian firms are rapidly adopting AI, but Indian firms also face infrastructure constraints. Currently, Indian data centers provide up to 1.2 gigawatts (GW) of operational capacity (Mandi et al., 2024).[4]

Policies that promote "compute" in India

Indian policies regarding the promotion of "compute" can be broadly divided into two: policies relating to computing infrastructure, and policies relating to development of semiconductors.

Policies on computing infrastructure

When it comes to supercomputers, the National Supercomputing Mission (NSM) was commenced in 2015 to "provide the country with supercomputing infrastructure to meet the increasing computational demands of academia, researchers, MSMEs, and startups by creating the capability design, manufacturing, of supercomputers indigenously in India" (Press Information Bureau of India, 2020). It would do so by installing 70 high-performance computing facilities at various research institutes and universities with a budget outlay of INR 4500 crores spread across 7 years (Press Information Bureau of India, 2015). The NSM endorses a variety of applications vital for national progress, such as climate modelling, weather forecasting, aerospace engineering, computational biology, disaster management simulations, and national security.

By the end of 2024, the government had installed 33 systems, which include the Param *Airawat, Arka*, and other systems mentioned earlier (Lok Sabha, 2024). The term of the NSM has been extended up to December 2025. The Indian government has considerable experience in the assembly and operation of supercomputers, so we expect the NSM to continue or be replaced with a similar but improved iteration after 2025.

When it comes to AI specifically, the government of India launched the *IndiaAI* Mission in March 2024. The government committed to spend INR 10,300 crores over the next five years to "ensure a structured implementation of the India AI Mission through a public-private partnership model aimed at nurturing India's AI innovation ecosystem" (Press Information Bureau of India, 2024a) The mission contains seven pillars. The first three — "compute" capacity, datasets, and application development — directly correlate with the AI "triad". The other pillars include startup financing, skill building, innovation funding, and "safe and trusted" regulation of AI.

Under the "compute" pillar, the Mission seeks to increase India's "compute" capacity by erecting "a cutting-edge, scalable AI computing infrastructure by deploying over 10,000 GPUs through strategic public-private collaborations" (Press Information Bureau of India, 2024a). Accordingly *IndiaAI*, an independent business division (IBD) under the Digital India Corporation (a non-profit company under the Ministry of Electronics and IT), issued a Request for Proposals (RFP) to "empanel agencies to provide AI services on cloud". Select bidders from the private sector will be chosen to operate tens of thousands of GPUs (Ministry of Electronics and IT, Government of India, 2024). The purpose is to set up a set of systems that are capable of training models and running AI-related compute instances for researchers, micro, small and medium enterprises (MSMEs), startups and other agencies which have obtained permission to do so from the IndiaAI mission ("end users"). The tenders were closed on 28 November 2024 and the bids were opened on

22 January 2025 (Ministry of Electronics and IT, Government of India, 2025a). In February 2025, ten bidders were shortlisted and they would begin offering bandwidth from 18,693 GPUs to the end users as specified by the IndiaAI scheme (Ministry of Electronics and IT, Government of India, 2025d). The current set of end users include researchers, academics, students, startups, micro, small and medium enterprises (MSMEs), along with the union and state governments and their various undertakings (Ministry of Electronics and IT, Government of India, 2025c).

The RFP seeks to "empanel", and not "procure", GPU usage and other services from private vendors. The firms empanelled by the Ministry are required to provide "AI compute instances, network services, storage services, AI platform and other AI services" to "end users". The government does not own the GPUs outright; rather they facilitate a right to use the GPUs and other services for the end users. Public expenditures here are intended to subsidize the costs of using the GPUs (Ministry of Electronics and IT, Government of India, 2024).

Policies on semiconductors

The Government of India also sees the development of a domestic semiconductor manufacturing industry as key to securing India's future role in the HPC supply chain.

The India Semiconductor Mission (ISM), initiated in 2021 with a budget of INR 76,000 crores across 5 years, is intended to be "a groundbreaking effort aimed at transforming India into a worldwide center for semiconductor manufacturing and design" (Press Information Bureau of India, 2024c). Its goals involve (i) offering financial support for fabrication and assembly facilities, (ii) encouraging the development of domestic intellectual property, and (iii) promoting both national and international partnerships. Under the ISM, we note the following schemes:

1. **Modified scheme for setting up of semiconductor and display fabs**: Under this scheme, the Union government would subsidise up to 50% of upfront capital cost for a fab that produces chips of wafer size 300 nm or less, with upfront total investment of more than USD 2.5 billion and potential revenue of USD 1 billion. For display fabs, the Union government subsidises up to 50% of upfront capital cost for a fab that produces TFT LCD or AMOLED.
2. **Modified scheme for setting up of compound semiconductors/ silicon photonics/ sensors/ discrete semiconductors/ outsourced semiconductor assembly and test (OSAT) facilities**: Under this scheme, the Union government subsidises up to 50% of upfront capital costs for a facility which assembles chips, with upfront total investment of more than USD 6.5 million.
3. **Design-linked incentive scheme**: Under this scheme, the Union government provides 50% of "eligible expenditure" up to INR 1.5 billion, as well as providing incentives between 4–6% on net sales turnover over 5 years subject to a ceiling of INR 3 billion. "Eligible expenditure" includes manpower costs, CapEx, IPR, field trials, and certain production costs.

In addition to this, the National Policy on Electronics 2019 (NPE) aims to establish India as an international center for Electronics System Design and Manufacturing (ESDM). The policy aims to meet the increasing domestic demand for electronics while decreasing reliance on imports (Ministry of Electronics and IT, Government of India, 2019). The NPE includes ambitious goals, such as reaching a turnover of USD 400 billion in the electronics sector by 2025, facilitating the production of 1 billion mobile devices worth USD 190 billion (with 600 million earmarked for export), and generating employment opportunities for 1 crore individuals. Along with other financial subsidies, some of which have been covered earlier, the policy includes the creation of a Sovereign Patent Fund. This fund aims to acquire essential intellectual property in the ESDM sector to foster domestic innovation in areas like fabless chip design and medical devices.

The production of semiconductors in India is a sign of how the Indian economy is gradually climbing the value-addition scale. Producing large chips that are required for non-intensive uses like household appliances, etc. at scale will certainly help generate technical expertise *provided* India is able to do so at low costs. While this does not have a direct bearing on high-performance computing in general, it does enable the transfer and development of skills and know-how.

The impact of external factors on "compute"

In this section, we talk about the impact of recent events in the field of high-performance computing on the Indian industry. We do so by using two examples: the supply-side constraints imposed by the United States through its recently announced *Framework for Artificial Intelligence Diffusion*, and the demand-side developments relating to the launch of DeepSeek, a Chinese AI lab that has released a new large language model. We also examine whether Indian policies are well-positioned, in the face of these developments, to ensure the balance of interests in India.

Developments in the United States

The first development we talk about is the recent Interim Final Rule (IFR) issued by the United States Department of Commerce (2025) on 15 January 2025. This rule governs how many GPUS can US manufacturers export to countries on the basis of their classification into three categories: Tier 1 countries (which include most of the NATO member states and non-NATO allies like Australia and Japan), Tier 2 countries (which include most other countries including India), and Tier 3 countries (which includes countries like China, Russia, and North Korea) (Heim, 2025). These rules apply even to US technology companies that operate data centers outside the United States.

The intent of the rule is clear: it seeks to restrict AI technology diffusion by restricting imports of GPUs. The rule, in fact, specifically targets Chinese access to AI technologies. At the same time, they also have the effect of limiting the scope of deploying these technologies in India. A US exporter supplying more than 1700 H-100 equivalent chips per year to an individual company in a Tier 2 country like India has to seek permission from the US Department of Commerce, and 100,000 H-100 equivalent chips per year for each country.[5] There are also new reporting requirements for US companies to detect and report "unauthorized training" of models in Tier 2 and Tier 3 countries. Some of these restrictions (such as those on unauthorized training) would apply even if a company from a Tier 2 country is "renting" a set of servers in a Tier 1 country. Some proposals in the United States have sought to enforce a KYC regime and end-user certificate for any customer in a Tier 2 or Tier 3 country that seeks to import chips from a US manufacturer (Heim & Egan, 2023). The rule is expected to come into force on 15 May 2025 (United States Department of Commerce, 2025).

Developments in China

Despite these new and older restrictions imposed by the US and other countries, we have also seen how the Chinese company DeepSeek has released its large language model "R1" (DeepSeek, 2025). The model is reported to be 27 times cheaper than OpenAI's competitor, o1, and it forces us to think about whether large language models can be developed without the need for large-scale hardware (Ball, 2025).

Perspectives and responses from India

How should Indian policymakers analyze these developments? *Firstly*, we note that in India, most AI products are *applications* and not models. To put the Indian state of play in the global context, it would be useful to use Lehdonvirta et al. (2024)'s positioning of countries with significant prowess in AI into "Compute North" and "Compute South" countries. Compute North countries are "positioned to use their territorial jurisdiction to intervene in AI development at the point at which models are sent to their local public cloud regions for training. For instance, they could require algorithms and data sets to be audited and certified for compliance with their local rules before training is permitted to commence, shaping what kinds of AI systems can enter the global market". Compute South countries are better suited for "AI system deployment than for development". Using this classification, we could put India somewhere in the middle — a country that is ahead of its peers when it comes to the deployment of models and applications to be generated out of it, but is not quite ready to make meaningful changes to the direction of the field itself.

Secondly, we note that the inputs for cloud computing, as Microsoft Corporation (2023) noted in their annual report, depend on "the availability of permitted and buildable land, predictable energy, networking supplies, and servers ...". It is well known that these input costs are high in India (Shah, 2015). These concerns may affect the diffusion of AI technologies in India (Medianama, 2025). Therefore, the impact of the US policies should, for now, be restricted to the *deployment* of AI applications. The DeepSeek development shows us that India could, if necessary, build, train and deploy its own large language model (LLM) even with limited "compute" capacity (Medianama, 2025).

Two interesting responses have emerged to these developments from the government of India. The first and older response is that of the idea of "AI sovereignty" and the creation of a "National AI Stack". In their concept paper, the Department of Telecommunications, Government of India (2020) identifies the strategic challenges regarding AI adoption in India (e.g., issues concerning access to technology, availability of trained researchers, etc.) It goes on to suggest that a National AI Stack could be a viable answer towards ensuring that all the functional requirements of AI research and services are met in India. Many of its suggestions are being implemented by the government. At the same time, academic commentators as well as the authors of the Concept Paper itself recognize the problems of unintended and unforeseen consequences with data privacy, ethical AI practices, and the regulatory environment (Pandey, 2025). These reasons, along with the AI adoption problems in India as mentioned in earlier, inform the need for caution and adequate safeguards in AI development in India.

The newer response has been that of the Indian government's recent RFP to build foundational models in India (Ministry of Electronics and IT, Government of India, 2025b). The developers selected to build an Indian LLM using various products developed under the other pillars of the IndiaAi mission: the developers would have access to GPUs acquired under the "compute" pillar and they would obtain research funding and training data for the same.[6] This development marks a credible and serious response by the Indian government to strengthen its commitment as a serious AI power. At the same time, concerns have arisen over whether these products would run the risk of underutilization, and therefore, are premature (Reddy et al., 2025).

Policy options to support improvement in "compute"

The single largest contributor to the Indian economy, by gross value added, is the IT services sector. Indian services personnel either work for global firms, headquartered in advanced economies, either through GCCs or through arms-length contracts with Indian firms (Shah, 2025). India's strengths had always laid in its large pool of skilled IT workers and these strengths transcend well into the HPC space. Many computers, such as the IBM's X10, made it into the Top 500 rankings on the strength of their efficient algorithms that "implemented its benchmarks elegantly" (Shah, 2007). These algorithms were developed at IBM's research labs in India. This was the same year that the Eka Supercomputer also entered the Top 500 rankings.

Therefore, we suggest that policy recommendations should necessarily keep this engine humming. They should encourage domestic talent and innovative applications for AI while reducing governmental frictions. Our suggestions are as follows:

"Buy" or "rent" should be a commercial decision

We appreciate the Ministry of Electronics and IT's decision to procure "compute" capacity through private sector service providers. We note that the government reduce frictions for the purely commercial decision of a private party to "buy" or "rent" infrastructure/ server space/GPUs. While the decision of the government to encourage the public availability of GPUs is encouraging, it would be most useful as a positive externality that is enjoyed by the scientific community and the MSME sector. These GPUs would, however, not form the bulk of deployment of applications for AI. Private sector/ for-profit use of GPUs, would. Private firms face the question of whether to buy or rent GPU space (also called infrastructure-as-a-service or IaaS). While this is a purely commercial decision, the government could certainly help reduce the frictions.

Reforms in procurement

Building capacity in the field of public procurement is like exercising a muscle: the more one does it, the better one gets at it (Chitgupi & Thomas, 2023). It is heartening to see the government provide the private sector with the space and capital to improve its capacities. The policy case for procurement, as opposed to in-house manufacturing, is clear (Mashelkar et al., 2024). However, public procurement, especially when it comes to R&D, tends to be over-monitored. R&D is inherently risky, and the Indian public procurement system was not geared to address contracting for projects with higher chances of failure. To address these issues, the best solution is to ensure legal clarity and certainty. One way to start this process is for the relevant ministries to develop a set of instructive manuals for the procurement of AI. These manuals should take into account the recency of the technology, the maturity of the technology (i.e., their technology readiness level or TRL), the risks associated with their deployment, and the consequent flexibility that the contract terms should adopt. For example, cost-plus contracts, which are generally discouraged by the General Financial Rules, could be a viable form of payment in the case of R&D contracts. When it comes to procurement of R&D, a relook at bidding requirements such as the bidder having a certain number of years of experience, etc. would be useful. Research grants are a powerful method of funding R&D in AI and the government must consider suitably amending the General Financial Rules.

Foreign policy that ensures the continued supply of "compute"

It has been reported that the news of Indian firms helping Russian firms evade sanctions on chips could have undermined India's credibility as a reliable partner of the United States (Bloomberg, 2024). Given that the United States and its allies control the overwhelming majority of the supply chain for high-performance computing, it is best to find diplomatic avenues to improve India's status to obtain preferential treatment for supplies of GPUs and other advanced processors. This has been done before — for example, in 1984, the United States relaxed certain supercomputer export restrictions on India (Subramanian, 2021). India is highly dependent on semiconductor imports and specialized components. In an uncertain trade environment — when various countries are beginning to increase tariffs on important goods — it is crucial that the Government of India avoid any tariff rammifications on AI R&D. This would be done by negotiating reduced tariffs on essential AI hardware inputs through bilateral and multilateral trade agreements can lower production costs and encourage domestic manufacturing.

Strategic trade partnerships with the countries that are major players in AI and semiconductors can facilitate technology transfer, allowing Indian firms to access cutting-edge innovations and collaborate on R&D. Additionally, easing foreign direct investment (FDI) norms and including AI-specific provisions in trade deals can attract global tech giants to set up fabrication units and data centers in India. By integrating trade agreements with initiatives like the India Semiconductor Mission and Make in India, the country can create a robust ecosystem for AI-driven computational infrastructure, positioning itself as a competitive global hub. Recently India has signed Memorandums of Understanding with Singapore and the European Union (Press Information Bureau of India, 2024b). These agreements revolve around joint R&D and production efforts for semiconductors meant for the development and deployment of AI. Such bilateral and regional agreements can be leveraged to achieve India's ambition to become a global leader in AI hardware and computational infrastructure.

References

Ball, D. (2025, January 28). *What DeepSeek R1 means — and what it doesn't*. Lawfare. https://www.lawfaremedia.org/article/what-deepseek-r1-means-and-what-it-doesn-t

Bloomberg. (2024). *How a Mumbai drugmaker is helping Putin get Nvidia AI chips*. https://www.bloomberg.com/news/features/2024-10-27/russia-is-getting-nvidia-ai-chips-from-an-indian-pharma-company

Brynjolfsson, E., Rock, D., & Syverson, C. (2019). Artificial intelligence and the modern productivity paradox: A clash of expectations and statistics. In *The economics of artificial intelligence: An agenda* (pp. 23–57). University of Chicago Press. http://www.nber.org/books/agra-1

Buchanan, B. (2020). *The AI Triad and what it means for national security strategy*. https://doi.org/10.51593/20200021

Chitgupi, A., & Thomas, S. (2023). *Learning by doing for public procurement*. *22*. https://ideas.repec.org/p/anf/wpaper/22.html

Das, K., & Sagara, H. (2017). State and the IT Industry in India: an overview. *Economic and Political Weekly*, *52*(41).

DeepSeek. (2025). *DeepSeek-R1: incentivizing reasoning capability in LLMs via reinforcement learning*. https://arxiv.org/pdf/2501.12948

Department of Telecommunications, Government of India. (2020). *Indian Artificial Intelligence Stack*. https://www.tec.gov.in/pdf/Whatsnew/ARTIFICIAL\%20INTELLIGENCE\%20-\%20INDIAN\%20STACK.pdf

Heim, L. (2025). *Understanding the artificial intelligence diffusion framework*. RAND Corporation. https://www.rand.org/pubs/perspectives/PEA3776-1.html

Heim, L., & Egan, J. (2023). *Accessing controlled AI Chips via Infrastructure-as-a-Service (IaaS): implications for export controls.* https://cdn.governance.ai/Accessing_Controlled_AI_Chips_via_Infrastructure-as-a-Service.pdf

Hemmen, L., & Pal, S. (2024). *Where is Europe's AI workforce coming from? Immigration, emigration and transborder movement of AI talent*. https://www.stiftung-nv.de/publications/where-is-europes-ai-workforce-coming-from

Indian Institute of Science. (2023, November 22). *IISc researchers part of international team winning prestigious Gordon Bell Prize*. https://iisc.ac.in/events/iisc-researchers-part-of-international-team-winning-prestigious-gordon-bell-prize-3/

Jujjavarapu, G., Hickok, E., & Sinha, A. (2018). *AI and the manufacturing and services industry in India*. Center for Internet; Society. https://cis-india.org/internet-governance/files/AIManufacturingandServices_Report_02.pdf

Kelkar, V., & Shah, A. (2022). *In service of the republic: the art and science of economic policy*.

Kotasthane, P., & Manchi, A. (2023). *When the chips are down: A deep dive into a global crisis.* Bloomsbury.

Lehdonvirta, V., Wu, B., & Hawkins, Z. (2024). *Compute North vs. Compute South: the uneven possibilities of compute-based AI governance around the globe.* https://doi.org/10.31235/osf.io/8yp7z

Lok Sabha. (2024). *National supercomputing mission*. https://sansad.in/getFile/loksabhaquestions/annex/183/AU381_gwXRcX.pdf?source=pqals

Mandi, S., Varadarajan, S., Mangalapalli, K., Lall, T., Srinivas, M., Chilamkurthy, S., & Rajagopalan, S. (2024). *Micro is the new mega*. Open Cloud Compute. https://pplus.ai/micromega

Mashelkar, R. A. (2008). Indian science, technology, and society: The changing landscape. *Technology in Society*, *30*, 299–308.

Mashelkar, R. A., Shah, A., & Thomas, S. (2024). Rethinking innovation policy in India: Amplifying spillovers through contracting-out. In *Accounting Theory and Practice* (Vol. 1, p. 100004). https://doi.org/10.1016/j.accoth.2024.100004

McKinsey. (2024, October 29). *AI power: Expanding data center capacity to meet growing demand*. https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/ai-power-expanding-data-center-capacity-to-meet-growing-demand

Medianama. (2025). *You can't scale AI in India with high costs: Nilekani advocates Rs 1 inference*. https://www.medianama.com/2025/04/223-ai-in-india-nandan-nilekani-india-digital-infrastructure/

Microsoft Corporation. (2023). *Annual report pursuant to section 13 or 15(d) of the Securities Exchange Act of 1934.* https://www.sec.gov/Archives/edgar/data/789019/000095017023035122/msft-20230630.htm

Ministry of Electronics and IT, Government of India. (2019). *National policy on electronics*. https://www.meity.gov.in/writereaddata/files/eGazette_Notification_NPE\%202019_dated\%2025022019.pdf

Ministry of Electronics and IT, Government of India. (2024). *Inviting applications for empanelment of agencies for providing AI services on cloud*. https://www.meity.gov.in/tenders/empanelment-agencies-cloud-based-ai-services

Ministry of Electronics and IT, Government of India. (2025a, January 22). *MEITY accelerates empanelment of agencies to provide AI compute and services for the IndiaAI mission*. https://indiaai.gov.in/article/meity-accelerates-empanelment-of-agencies-to-provide-ai-compute-and-services-for-the-indiaai-mission

Ministry of Electronics and IT, Government of India. (2025b). *IndiaAI mission: call for proposals to build foundational AI models*. https://indiaai.gov.in/article/indiaai-mission-call-for-proposals-to-build-foundational-ai-models

Ministry of Electronics and IT, Government of India. (2025c). *End user policy for AI services on the cloud*. https://indiaai.s3.ap-south-1.amazonaws.com/docs/end-user-policy-for-indiaai-compute-portal.pdf

Ministry of Electronics and IT, Government of India. (2025d). *Regarding availability of AI compute, network, storage, and AI platform services at affordable rates*. https://indiaai.s3.ap-south-1.amazonaws.com/docs/office-memorandum-india.pdf

Mint. (2025). *MEITY picks Sarvam AI to lead India's sovereign LLM effort; Vaishnaw vows global competition*. https://www.livemint.com/technology/tech-news/india-taps-sarvam-ai-to-build-first-sovereign-llm-under-indiaai-mission-ashwini-vaishnaw-11745660295745.html

NASSCOM. (2024a). *AI powered tech services: a roadmap for future ready firms*. https://web-assets.bcg.com/f6/5a/abdb70be44749d78ef21fafc89c5/ai-powered-tech-services-a-roadmap-for-future-ready-firms.pdf

NASSCOM. (2024b). *GCCs in India: Building resilience for sustainable growth.* https://nasscom.in/system/files/publication/GCCs\%20in\%20India\%20\-\%20Building\%20resilience\%20for\%20sustainable\%20growth_Web\%20final.pdf

Organisation for Economic Co-operation and Development. (2019). *Recommendation of the Council on Artificial Intelligence* (OECD/LEGAL/0449). https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449

Organisation for Economic Co-operation and Development. (2021). *The human capital behind AI: Jobs and skills demand from online job postings*. *120*. https://doi.org/10.1787/2e278150-en

Organisation for Economic Co-operation and Development. (2023). *A blueprint for building national compute capacity for artificial intelligence* (350). https://doi.org/10.1787/876367e3-en

Pandey, P. (2025). Digital sovereignty and AI: developing India's National AI Stack for strategic autonomy. *Procedia Computer Science*, *254*, 250–259. https://doi.org/10.1016/j.procs.2025.02.084

Press Information Bureau of India. (2015). *National Supercomputing Mission (NSM): Building Capacity and Capability*. https://pib.gov.in/newsite/printrelease.aspx?relid=117702

Press Information Bureau of India. (2020). *National Supercomputing Mission: a transformative approach in supercomputing*. https://pib.gov.in/newsite/PrintRelease.aspx?relid=200608

Press Information Bureau of India. (2024a). *Cabinet approves over Rs 10,300 crore for IndiaAI mission, will empower AI startups and expand compute infrastructure access*. https://pib.gov.in/PressReleasePage.aspx?PRID=2012375

Press Information Bureau of India. (2024b, September 5). *Prime Minister's meeting with Prime Minister of Singapore*. https://pib.gov.in/PressReleasePage.aspx?PRID=2052028

Press Information Bureau of India. (2024c, November 17). *Assam's semiconductor plant: a game-changer for India's semiconductor ecosystem*. https://pib.gov.in/PressReleaseIframePage.aspx?PRID=2074074

Reddy, B., Panicker, R., Krishna, S., Goswami, A., Sen, A., & Madhavan, A. (2025). *State of AI governance, 2024*. Takshashila Institution. https://static1.squarespace.com/static/618a55c4cb03246776b68559/t/68022f8c5e058d397dcea587/1744973717444/SOAIG.pdf

Sanglard, F. (2020, May 2). *A history of NVidia Stream Multiprocessors*. https://fabiensanglard.net/cuda/index.html

Sastry, G., Heim, L., Belfield, H., Anderljung, M., Brundage, M., Hazell, J., O'Keefe, C., Hadfield, G. K., Ngo, R., Pilz, K., Gor, G., Bluemke, E., Shoker, S., Egan, J., Trager, R. F., Avin, S., Weller, A., Bengio, Y., & Coyle, D. (2024). *Computing power and the governance of artificial intelligence*. https://arxiv.org/abs/2402.08797

Shah, A. (2007). Interpreting the 4th fastest supercomputer in the world. *The Leap Blog*. https://blog.theleapjournal.org/2007/11/interpreting-4th-fastest-supercomputer.html

Shah, A. (2015, March 15). *The economics of cloud computing: an Indian perspective*. https://blog.theleapjournal.org/2015/03/the-economics-of-cloud-computing-indian.html

Shah, A. (2022). The industry structure of India's large firms: IT is the biggest industry. In *The Leap Blog*. https://blog.theleapjournal.org/2022/03/the-industry-structure-of-indias-large.html

Shah, A. (2025, January 20). *US export restrictions on AI chips: Implications for India*. Business Standard. https://www.mayin.org/ajayshah/MEDIA/2025/ai_chips_implications.html

Singh, N. (2016). Information technology and its role in India's economic development: a review. In *Development in India*. Springer. https://doi.org/10.1007/978-81-322-2541-6_14

Sterling, T., Brodowicz, M., & Anderson, M. (2017). *High performance computing: Modern systems and practices*. Morgan Kaufmann.
Subramanian, R. (2021, November 19). *Monsoons, computers, satellites: History and politics of weather monitoring in India*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4021101

United States Department of Commerce. (2023). *Implementation of additional export controls: Certain advanced computing items; supercomputer and semiconductor end use; updates and corrections*. https://www.bis.doc.gov/index.php/documents/federal-register-notices-1/3353-2023-10-16-advanced-computing-supercomputing-ifr/file

United States Department of Commerce. (2025). *Framework for artificial intelligence diffusion*. https://www.govinfo.gov/content/pkg/FR-2025-01-15/pdf/2025-00636.pdf

Vipra, J., & West, S. M. (2023). *Computational power and AI*. https://ainowinstitute.org/publication/policy/compute-and-ai#3e984e81-6a26-4d11-8ed6-793481718cb0

Wolcott, P., Homer, P., & Goodman, S. (1998). HPC export controls: navigating choppy waters. *Communications of the ACM*, *41*(11), 27–30. https://doi.org/10.1145/287831.287836

Appendix

Industry	% of total non-financial GVA		
	2013	2018	2023
Information technology	16.8	22.75	25.51
Chemicals & chemical products	13.79	16.86	15.69
Transport equipment	9.69	10.05	7.8
Crude oil & natural gas	9.15	4.36	3.35
Metals & metal products	8.36	5.48	7.73
Electricity generation	5.17	4.07	3.79
Industrial & infrastructural construction	4.92	3.62	3.67
Communication services	4.54	0.99	2.09
Food & agro-based products	3.28	3.23	3.4
Machinery	3.12	3.11	2.33
Wholesale & retail trading	2.71	3.71	4.08
Construction materials	2.52	2.41	2.07
Coal & lignite	2.27	2.1	1.54
Textiles	2.26	1.01	1.83
Miscellaneous services	1.81	4.91	5.43
Consumer goods	1.61	2.4	2.14
Transport services	1.33	2	1.52
Diversified manufacturing	1.18	0.53	0.41
Minerals	1.07	0.53	0.65
Electricity transmission & distribution	1.05	1.9	1.18
Diversified	0.9	0.78	1.14
Diversified non-financial services	0.74	0.27	0.87
Real estate	0.71	0.91	0.48
Miscellaneous manufacturing	0.53	0.97	0.87
Hotels & tourism	0.47	1.06	0.41

Percentage contribution to gross value added (GVA) among all non-financial sectors. A crude measure of GVA used here is the sum of profits before tax, depreciation, and wages. Source: CMIE Prowess (2013–2023).

Endnotes

1 "Machine learning" refers to AI systems which are "able to learn from data and to identify patterns to make decisions with minimal human intervention". See Organisation for Economic Co-operation and Development (2021).

2 "Invisible infrastructure" denotes the "hard" infrastructure of GPUs and servers, and the "soft" infrastructure of the state. For further reading, see Kelkar & Shah (2022).

3 US Section 15 CFR 772(i), see United States Department of Commerce (2023).

4 For the sake of comparison, the United States has over 55 GW, the United Kingdom has around 2.5 GW, see McKinsey (2024).

5 Each H-100 equivalent can deliver between 6–8 PFlops.

6 Recent media reports indicate that Sarvam AI, a Bangalore-based AI service provider, has been selected to develop the LLM Mint (2025).

PAHLE INDIA FOUNDATION

{X}KDR

AI KNOWLEDGE CONSORTIUM