FJWU (Economics) Working Papers
WP2026:1

The Compute Supercycle: Infrastructure, Geo-politics and the Structural Transformation of Global Growth

Jyotsna Thakur
Shyamsundar Subramani
Arshad Bhat

Department of Economics
Fatima Jinnah Women University

FJWU (Economics) Working Papers
WP2026:1

The Compute Supercycle: Infrastructure, Geo-politics and the Structural Transformation of Global Growth

Jyotsna Thakur
Shyamsundar Subramani
Amity School of Engineering and Technology
Arshad Bhat
Amity Institute of Liberal Arts
Amity University Mumbai, Maharashtra, India

Department of Economics
Fatima Jinnah Women University
2026

Editorial Committee

Dr. Bushra Yasmin

Dr. Saira Tufail

Dr. Sadia Sherbaz

Disclaimer:

The Department of Economics, *FJWU* Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the author (s). Copyrights to this FJWU Economics Working Paper remain with the author (s). The author (s) may publish the paper, in part or whole, in any journal of their choice.

Recommended Citation :

Thakur, J., Subramani, S., & Bhat, A. (2026). The Compute Supercycle: Infrastructure, Geopolitics and the Structural Transformation of Global Growth. FJWU Economics Working Paper Series (WP2026:1). Fatima Jinnah Women University. Rawalpindi.

Department of Economics, Fatima Jinnah Women University,

The Mall, Rawalpindi

Email: ecowps@fjwu.edu.pk

Website: http://fjwu.edu.pk

Contact: 051-9292900, Ext: 2072

The Compute Supercycle: Infrastructure, Geopolitics and the Structural Transformation of Global Growth

Jyotsna Thakur, Shyamsundar Subramani & Arshad Bhat

Abstract

From how AI hardware and energy systems intersect, to the global economy's shift toward more resource-intensive servers and low-scale computing, this paper argues that this "compute supercycle" is not just a trend in tech spending, but a reinvention of the global economy. We create a single model, compute-economics, that brings together three blocks: AI capital expenditure, semiconductors supply chain, data-center energy usage and project, region and industry specific AI productivity gains, leveraging primary data from IMF, World Bank, IEA, WSTS, IDC, FRED, Goldman Sachs, Morgan Stanley and World Economic Forum. On the practical front, we show that in 2026 annual AI CAPEX will be 1.4 trillion, with total capacity growth in data centers expected to be 847 GW and the revenue from semiconductors will be 912 billion in 2026.Empirically, for instance, you can see that in 2026 the annual investment in AI will be 1.4 trillion, the power consumption of data centers will be 847 GW, and the AI's semiconductors revenue will be 912 billion. Extreme concentration of fabrication in Taiwan, energy constraints degreeed as more and more data centers drive electricity consumption, and the more import of advanced chips are controlled – three risk factors get studied and we try to gauge their likely effect on both bull, base and bear scenarios for compute demand, semiconductors and energy infrastructure buildup. Methodologically, we introduce the Compute Economics 2026 Dashboard, a multi tab visualization and analytics platform that is real time and becomes the platform for operationalising this framework for investors, policy makers and corporate strategists; bringing different types of heterogeneous macroeconomic, market and technical indicators together in one decision interface.

Keywords: *Compute; Semiconductors; Infrastructure; Geopolitics; Growth*

JEL Codes: *O33, O14, O41, L63, Q43, F52*

1. Introduction

In the 2026 world economy, what analysts call the "compute supercycle", the round of massive investments in computing hardware to support the "industrialization of Generative Artificial Intelligence", is driving in most part the world. Contrary to previous technology cycles focused on "soft" elasticity of software components and virtualization of clouds, the current shift in technology seems to be motivated by "hard" constraints and requirements of compute infrastructure: through-put-density of semiconductors, massive energy purchases and strategic issues of international supply chain management [4][5][6].

The investment being made in this supercycle is unprecedented – the global market for investments in AI is projected to reach approximately $1.4 trillion a year by 2026 [2] [3] [5]. The fast rollout is revolutionizing digital and physical worlds and is expected to consume 4.2 per cent of the world's electricity, more than any other technology, as AI evolves from being treated as a trial to becoming a key embedding of technology as infrastructure. So much so that great power silicon and energy reliability and dispatchability have become a vital asset alongside labour or traditional capital to the economic growth path [7][8].

The development of this infrastructure at such a fast pace has brought to the fore threats which are influencing geopolitical and economic dynamics in 2026 [9]. The intense focus of the advanced-node semiconductor industry in geopolitically sensitive nations, an increasingly stringent international controls framework, and emerging energy constraints are complex factors that result in infrastructure availability as the limiting factor for growth-besides capital-on top of the semiconductor market. The components found on top of the semiconductor market are of geopolitical concern, an increasingly stringent international controls regime and emerging energy constraints, but likely most importantly, not only capital, but infrastructure is the limiting factor for advancing into the semiconductor market at advanced nodes. This paper examines these interdependent drivers, to develop a comprehensive picture of the impact of compute-centric intelligence, the energy market's transformation and state-led technology policy as they affect productivity and economic competition in the modern world.

2. Review of Literature

Current academic and industry discussion around "compute economics" and the AI supercycle has turned from prior software-defined technology cycles to change the focus to the physical underpinnings of Artificial Intelligence (AI) applications [1][2].

2.1 The AI Infrastructure Supercycle

Current research has focused beyond pioneers to a cycle of innovation, focusing not only on innovation, but also on a "multi decade" structural capital expansion model focused on physical infrastructure [3][5][7]. This supercycle will focus on the continued capitalising of "real" infrastructure, in contrast with the "cloud-centric" models of previous supercycles. The analysts note that the "physicalization" of intelligence-that is, control of massive compute clusters and power generation capacity-is the new chief driver of the firm-level and macro-economic productivity. This change represents a new time of measurement systems that are no longer software elasticity in terms of economic scale and valuation, but the throughput of semiconductor devices and raw FLOPs per watt.

2.2 Semiconductor Supply Chain Dynamics

The debate on semiconductor supply chains is largely one of the clashes of history's globalization and technological decoiling [14]. Fabrication of advanced nodes (Under 5nm) is another systematic deficiency that is pointed out by research more and more with the centralization. Geopolitical risks in such dependencies are actively being modeled and the CHIPS Acts and other measures via the state are driving towards strategic autonomy [15][16]. According to the book, the focus has shifted from optimising supplies with priorities and optimization from supply chain efficiencies to building parallel technological ecosystems, or secundary value chains, that can help mitigate against unexpected downtimes in trading supply chains.

2.3 Energy Constraints and Grid Integration

The literature on the issue of AI and Energy has talked about a rapidly building problem — because these data centers are becoming large, localized power consumers all over the world. AI consumes electricity, and the forecasted consumption of AI frontier models is surpassing traditional energy efficiency gains [18][19] which necessitates the introduction of dispatchable baseload energy. One of the key points in the research cited here is the need for AI training jobs to support high uptime, which often conflicts with the intermittent nature of renewable energy resources. The need for large-scale investments in nuclear power, energy storage, and grid scaling to avoid stifling compute scaling with energy shortages is now increasingly coming to the fore.

2.4 Economic and Labour Market Implications

An evaluation of whether AI might herald a structural regime shift in total factor productivity (TFP) as compared to the period following the 2008 financial crisis is underway in the economic literature [20]. Growing research agendas are split on the effects of the automation of routine cognitive tasks on labour markets, with some research focussing on the danger of further automation rendering low-skill jobs redundant, while other studies examine the potential of new and augmentative employment categories emerging in response to the introduction of new technologies. In addition, the "democratization of compute" is being studied and exemplified in recent works of cloud-based AI infrastructure, indicating the possibility of distributing the gains of AI productivity to a new population of emerging markets unavailable to access the same boon due to their lack of a full local semiconductor and hardware ecosystem.

3. Methodology and System Architecture

In this study, data-synthesis research is used to analyze the interaction between the four dimensions – Compute demand and semiconductor production, Energy infrastructure, and Geopolitical risk – with economic performance of the world in the year of 2026 utilizing mixed-method research. The methodology does not limit to describing the reporting – it gives a highly integrated (meconomic) framework of – macroeconomic indicators – sectoral market indicators – measures of infrastructure features – geopolitical values and expands the basis for calculation econometrics-model information. The system architecture employs the Compute Economics 2026 Dashboard, which is a multi-layer interactive application that processes, harmonizes and analyzes on the fly all real-time and periodic data from various sources of institutional, market and policy data, while visualizing the data.

3.1 Research Design

The research uses a multi-source intelligence synthesis approach to secondary data. The strategy here was to gather structured and semi-structured data from officially published finance and macroeconomic institutions, from data sources representing the semiconductor industry, official industry data and energy industry databases, market feeds, and public policy databases, and to normalize the datasets into a common data analysis framework for comparative analysis. The goal is to go beyond 'sector analysis' and instead considers the interdependencies between compute investment, hardware supply, electricity demand, as well as strategic regulation.

Four analytical dimensions – compute expansion, semiconductor market structure, energy-system stress and geopolitical control mechanisms – have been organized around the study. The dimensions are now finally connected to national and global growth indicators, using an interpretive dashboard model that facilitates growth monitoring, forecasting and generation of insights for the planning process.

Table 1. Core Methodological Framework

Component	Method Used	Purpose	Output
Macroeconomic analysis	Secondary data synthesis	Measure AI contribution to GDP and growth	GDP impact indicators

Semiconductor analysis	Market concentration and capacity mapping	Assess supply-side constraints and vendor dominance	Fab capacity and market share analysis
Energy analysis	Infrastructure demand correlation	Track AI-related electricity demand and energy mix	Energy demand and pricing relationship
Geopolitical analysis	Risk scoring and policy tracking	Evaluate export controls and supply chain vulnerabilities	Trade-risk and policy-risk indicators
Forecasting	Scenario-based comparative modeling	Estimate 2026–2030 compute trajectories	Bull, base, and bear scenarios
Visualization	Interactive dashboard analytics	Convert data into decision-ready intelligence	Real-time multi-tab interface

3.2 Data Collection Strategy

Several authorities of data of area are applied in the collection of information for the study, as to acquire a wide area, comparability and dependability. The methodology focuses on and puts great emphasis on institutional data and data sources that have industry expertise, including IMF, World Bank, IEA, WSTS, IDC, FRED, BIS, Goldman Sachs and Morgan Stanley. Every source provides its own facet to the compute economy, including forecasts of GDP growth, revenues on semiconductors, deployments of GPUs, energy usage, energy prices, and export-control events.

To mitigate the source bias, cross domain triangulation is adopted in the study, in which no sole indicator is used for interpretation. For instance, we look at capital spending of AI and compare it with semiconductors revenue, data center power usage, and GDP contribution to figure if capital flows are coming from real growth in infrastructure and macroeconomics.

Table 2. Data Sources Used in the Study

Source	Data Category	Frequency	Method of Use
IMF WEO	GDP forecasts, AI growth contribution	Quarterly	Macroeconomic modeling
World Bank	Emerging market investment, development indicators	Quarterly	Regional investment analysis
IEA	Data-center power demand, energy mix	Monthly	Energy-system analysis
WSTS	Semiconductor revenue, manufacturing trends	Monthly	Chip-market analysis
IDC	GPU shipments, deployment estimates	Quarterly	Compute expansion analysis
FRED	Natural gas and electricity price indicators	Daily	Energy-price correlation

BIS	Export controls, restrictions, entity actions	Event-based	Geopolitical policy analysis
Goldman Sachs / Morgan Stanley	AI capex, scenario forecasts, risk outlook	Weekly / Monthly	Forward-looking market analysis

3.3 Data Processing and Normalization

The collected datasets have different units of measurement, frequency, geographic scope and purpose of analysis, so a normalisation process comes before them being inserted into the dashboard model. Monetary values are shown in U.S. dollars; growth values have been normalized, if appropriate, to show the year-over-year percent change; forecast categories are distinguished; and actual, estimate, and consensus projections have been clearly separated. Time-series data are restructured in a comparable manner to enable a comparison of energy, semiconductor and macroeconomic indicators for the same time frame.

The system also has rules for traceability, which also help preserve the source transparency. Each quantitative metric is associated with a named source and an established update cycle (e.g., daily, weekly, monthly, etc.), and out-of-date data more than twice the expected update interval are marked as stale data to avoid misinterpreted data analysis. This approach is also suitable for policy and investment monitoring in real-time application.

3.4 Analytical Model Construction

The primary methodological development of the study is the formulation of an analytically tractable model to connect infrastructure traits with macroeconomic and geopolitic consequences. A remarkable methodological achievement of this study is that the model becomes analytically manageable by linking infrastructure traits with macroeconomic and geopolitical consequences. The model differs in conceptualizing AI as a software phenomenon, instead considering compute as a strategic production input influenced by several factors, including hardware availability, energy capacity, trade exposure and state policy. It is therefore presented as a pipelined model with input and transformations, risk-outputs, and strategic-outputs.

There are number of variables at the input level in the model, for example, the cost of AI products and services, deployment of GPUs, semiconductors' capacity, electricity demand, energy prices, policy constraints, and so forth. All of these inputs get fed into the system and used in generated comparative indexes such as Compute Growth Index, supply chain risk score, GDP contribution estimates and scenario forecasts. The results of these processes are then presented in a visual way using visual modules, so that the results of the process can be easily understood by investors, analysts and strategic decision makers.

3.5 Scenario Modeling Method

The study employs scenario modelling to make estimates about possible futures from 2026 to 2030. There are three scenarios: Bull, Base and Bear. The scenarios vary, depending on assumptions about the rate of progress in A.I. capabilities, how strong semiconductors will be next year and beyond, how fast energy will be produced and how much geopolitical turmoil will exist, particularly regarding the risks in Taiwan and restrictions on exports.

The bull scenario is based on the following assumptions: fast capability growth, rapid infrastructure scaling, and no bottlenecks. In the base scenario it is expected that the expansion takes place in a steady manner but is limited by the smaller amount of friction. Bear scenario: Postulates a significant geopolitical conflict, regulatory shock, or energy constraints that constrain compute growth, and ultimately reduce economic returns.

Table 3. Scenario Modeling Assumptions

Scenario	Probability	Core Assumption	Expected Outcome
Bull	18%	Rapid AI progress, energy bottlenecks relieved, stable geopolitics	Very high compute and GDP expansion
Base	55%	Strong but moderated growth with partial infrastructure constraints	Sustained supercycle continuation
Bear	27%	Major semiconductor, policy, or grid disruption	Capex slowdown and reduced GDP gains

3.6 Functional Module Design

It's divided into 7 analytical modules: Overview, AI Chips, Energy, Geopolitics, Global Growth, Forecasts and Insights. All modules are aligned to a larger dimension of research, and linked to chart types, KPI cards, and narrative outputs. Importantly, this modular design keeps two values independent of each other, with one being an observation of the facts and the other being an interpretation, yet at the same time does not obscure the interconnection between variables.

Overview includes KPI cards, a regional capex overview chart, a compute growth line and a 3D globe for a macro overview. The AI Chips, Energy and Geopolitics modules add to sectoral analysis, and the Global Growth and Forecasts link infrastructure change and its impact on macro-economic projections with future scenarios. The Insights module then pinpoints key anecdotal patterns and trends based on the aggregated evidence in the form of discs and a confidence matrix.

4. Results

The findings highlight that the compute economy is not simply a technology-sector phenomenon anymore, but rather is a sweeping structural shift that is realized through transistors--whether semiconductor-based or not--as well as energy systems and macroeconomic growth across the economy. Considering all the indicators as a whole, the most consistent trend is that capital expenditure, hardware spending and electricity consumption growth all accelerated at the same time as well as GDP contributions, in line with the paper's thesis that the compute supercycle is not simply a passing fad, but represents a structural shift.

4.1 Aggregate Compute Expansion

The first big find is scale and speed of computing growth predicted by 2026.The first big discovery is the magnitude and the velocity of compute expansion by 2026. Global capital expenditure for AI grew by nearly 50% year-on-year to 1.4 trillion, with computer output increasing by over 100% to 4.8 zettawatts, global demand for power in data centers more than doubled to 847 GW and total revenue from semiconductors more than doubled to 912 billion, with all of the above metrics seeing significant increases over the previous year. The Compute Growth Index (CGI) also moved up, from a baseline of 100 in 2020 to 820 in 2026, which equates to 8.2x times growth in the six years.

Such results indicate that the development of the AI infrastructure is not just an overvaluation dream by itself. But measurable gains in physical deployment, market revenue and energy use provides support for increasing industrialization of AI systems, rather than for "compute.

Table 4. Key Compute Economy Indicators in 2026

Indicator	2026 Value	YoY Change	Interpretation
Global AI Capex	1.4T	38%	Strong capital deepening in AI infrastructure
Compute Output	4.8 ZW	61%	Rapid increase in usable AI compute capacity
Data Center Power Demand	847 GW	29%	Infrastructure growth is energy intensive
Semiconductor Revenue	912B	22%	Chip markets are expanding with AI demand
GPUs Deployed	3.2B units	44%	Hardware adoption continues at scale
Compute Growth Index	820	From 100 in 2020	Long-cycle structural acceleration

4.3 Top Conferences for AI Developers

The regional results show that the distribution of AI investment in capital expenditures is still quite regionalised. North America takes 52 percent of the global share of AI capex, the Middle East 5 percent, Asia-Pacific 12 percent and Europe 3 percent of AI capex, while the remaining 3 percent comes from the rest of the world. This focus documents current strengths in the hyperscaler investment, the semiconductor ecosystem, energy access and policy coordination.

The results reflect at the same time the start of geographic diversification. Investment in AI in emerging markets - any nation outside of the traditional U.S.-East Asia investment spectrum - jumped 68 percent year over year to $184 billion in 2026.

4.3 Semiconductor Market Results

Semiconductor findings show that the semiconductor industry is very concentrated with NVIDIA and TSMC and only a handful of processors, memory and packages. In the open-market share, NVIDIA accounted for 78% share of AI GPUs while AMD accounted for 10%, Google TPU 5%, Intel Gaudi 3% and others 4%. In addition to its dominance, NVIDIA's AI revenue estimate of 248 billion was also early signs of a custom silicon alternative as both competitors, AMD and Broadcom (with its custom XPU business), delivered 31 billion and 48 billion, respectively.

The other thing that emerged in the last few semiconductors is that there's a high number of manufacturing facilities all concentrated right in East Asia. Taiwanese capacity makes up 68 percent of the world's sub-5nm capacity, South Korea's 18 percent, the United States' 4 percent, Japan's 2 percent, China's 2 percent and Europe's 3 percent. The number of these earmarks just confirms that with the advanced AI compute continues to be based on a narrow geographical manufacturing footprint.

.**Table 5. AI GPU Vendor Market Share, 2026**

Vendor	Market Share	Result Interpretation
NVIDIA	78%	Clear platform and ecosystem dominance
AMD	10%	Secondary challenger gaining selectively
Google TPU	5%	Important but mostly internal deployment
Intel Gaudi	3%	Limited competitive traction
Others	4%	Fragmented residual market

Table 6. Advanced Semiconductor Fabrication Capacity by Region, sub-5nm

Region/Country	Share of Global sub-5nm Capacity	Implication
Taiwan	68%	Highest systemic concentration risk
South Korea	18%	Secondary advanced-node center
United States	4%	Rebuilding but still limited capacity
Europe	3%	Early-stage strategic catch-up
Japan	2%	Limited but important re-entry
China	2%	Constrained by controls and tooling limits

4.4 Energy System Results

The findings regarding energy show that AI infrastructure expansion is now influencing electricity grids considerably. Data centers consumed 4.2 percent of all electricity in the world in 2026, an increase compared to 2.4 percent two years ago, while 41 percent of the power supply for AI data centers came from renewable energy sources, and 38 GW of nuclear power capacity has already been allocated for AI requirements. The largest single energy source for AI infrastructure was natural gas at 38 percent, followed by solar energy at 18 percent, wind energy at 14 percent, coal at 14 percent, nuclear energy at 9 percent, and hydropower at 7 percent.

The demand for electricity on a country level is also highly irregular. The US was the maximum electricity consumer for AI data centers at 320 GW, while China was the second largest user at 218 GW, the EU was in third place with 153 GW, Japan occupied fourth place at 77 GW, India was in fifth place at 62 GW, and the UAE–Saudi Arabia corridor was in the sixth place at 44 GW. This indicates that the supremacy in computing increasingly depends on the strength of electric grids.

Table 7. Energy Profile of the AI Economy, 2026

Metric	Value	Result Interpretation
Data center share of global electricity	4.2%	AI is becoming a macro-energy issue
Renewable share in AI data centers	41%	Decarbonization is progressing but incomplete
Nuclear capacity committed	38 GW	Baseload demand is reshaping energy planning
Natural gas price	3.84/MMBtu	Gas remains critical to flexibility and reliability

4.5 Geopolitical Risk Results

The consequences have shown that the importance of semiconductor supply chains lies in their much-exposed degree of concentration and risk from global policies. The scores for the different degrees of concentration achieved very high levels, with Taiwan's concentration score being 9.2, while the advanced packaging risk and rare earth exposure to China gained 8.2 and 7.8 respectively, ASML EUV access scored 7.5, while the concentration of HBM memory came to 6.5. The U.S. domestic fabs' degree of concentration was the lowest at only 3.8, which is an improvement, but not enough for achieving independence in this industry.

There is also an observed increase in the intensity of export controls. The United States leads with 84 current restrictions, while the Netherlands follows with only 42, Japan had 38, the United Kingdom had 28, the European Union's score was 24, Korea's score was at 18, Australia at 14, and Canada at 12. The results show that the growing competition in computing technologies is now defined not by the market but by industrial policy and technology restriction.

Table 8. Semiconductor Supply Chain Risk Scores, 2026

Supply Chain Node	Risk Score (0–10)	Risk Level	Main Concern
Taiwan advanced nodes	9.2	Critical	Extreme concentration of leading-edge production
Advanced packaging	8.2	Very High	Limited supplier and location diversity
Rare earths (China)	7.8	High	Material export dependency
ASML EUV access	7.5	High	Single-source equipment vulnerability
HBM memory	6.5	Elevated	Concentrated high-bandwidth memory supply
U.S. domestic fabs	3.8	Manageable	Capacity improving but still limited

4.6 Global Growth and Productivity Results

The macroeconomic data demonstrate that AI is already making a significant contribution to the development of national and the global economy. According to the estimation based on IMF data presented in this paper, the contribution of AI to global GDP growth in 2026 is expected to reach 1.8 percentage points, while global GDP is anticipated to grow by 3.4 percent, and total factor productivity of the most advanced economies is expected to be 14 percent. Moreover, this information proves that AI is transforming into an input factor contributing to productivity growth and not just a technology applicable to a particular industry. The comparison across the countries reveals that the outcomes of growth differ tremendously. On the list of countries experiencing the largest increase in GDP caused by AI, the United Arab Emirates are on top with an estimated increase of 3.1 percentage points, whereas India comes second with 2.4 percentage points of GDP increase due to AI; the United States possesses 2.1 percentage points, China – 1.9 percent, the United Kingdom – 1.7 percentage points, Canada – 1.6 percentage points, the European Union as a whole is estimated to have 1.4 percentage point increase due to AI; Germany has 1.3 percentage points of increase, Brazil – 1.2; Japan is estimated to experience 1.1 percentage growth due to AI

Table 9. AI-Driven GDP Impact by Country/Region, 2026E

Country/Region	AI GDP Boost	AI Investment	Key Result
UAE	3.1 pp	38B	Highest relative GDP gain
India	2.4 pp	64B	Strong growth with lower capital base
United States	2.1 pp	480B	Largest absolute compute investment
China	1.9 pp	320B	Strong scale under policy constraints
United Kingdom	1.7 pp	NA	Service-sector productivity effect
Canada	1.6 pp	NA	Talent and commercialization advantage
European Union	1.4 pp	210B	Large market but slower diffusion
Germany	1.3 pp	NA	Industrial optimization-led gains

Brazil	1.2 pp	NA	Emerging-market AI application growth
Japan	1.1 pp	88B	Automation and robotics-driven gains

Table 10. Sector Productivity Gains from AI Deployment

Rank	Sector	Productivity Gain	Main Use Cases
1	Financial Services	31%	Trading, fraud detection, risk modeling
2	Healthcare & Life Sciences	27%	Drug discovery, imaging, trials
3	Software & Technology	25%	Code generation, testing, DevOps
4	Legal & Professional Services	22%	Review, research, compliance
5	Manufacturing	18%	Predictive maintenance, quality control
6	Retail & Logistics	14%	Forecasting, routing, merchandising

4.7 Forecast Results

The forecasting results indicate that future growth will continue to occur across the different scenarios, although the extent may differ widely based on the expansion of energy infrastructure and stability of geopolitical affairs. In the high-case scenario, demand for AI computing would reach 82 zettaFLOPs by 2030, as compared to the base-case forecast of 28 zettaFLOPs and the low-case expectation of 14 zettaFLOPs. The revenues in the semiconductor industry are expected to rise from 912 billion in 2026 to 1.82 trillion in 2030 whereas the demand for AI data center electricity is projected to grow from 847 to 2,100 GW.

The analysis suggests that the future development will not depend on technological improvements alone but rather on the effectiveness of electricity development and energy diversification as well as changing trends in geopolitical affairs.

Table 11. Forecast Trend of Core Compute Metrics, 2026–2030

Metric	2026E	2027E	2028E	2029E	2030E
Semiconductor Revenue (USD B)	912	1,080	1,280	1,540	1,820
AI DC Power Demand (GW)	847	1,100	1,380	1,720	2,100
AI GDP Contribution (pp)	1.8	—	2.8	—	4.2
GPU Shipments (M units)	38.0	—	82.0	—	160.0

The figure 1 shows dashboard that summarizes rapid global expansion of AI compute, capex, and semiconductor activity from 2020–2026, including regional investment shares and associated growth indicators.

Figure 1. Compute Economics 2026 Dashboard Showing Global AI Infrastructure, Investment, and Growth Indicators (2020–2026).

The figure 2 shows a highly concentrated global AI chip landscape, with AI GPU market share and advanced-node fab capacity dominated by a few vendors and regions, alongside rapidly rising AI vendor and semiconductor revenues.

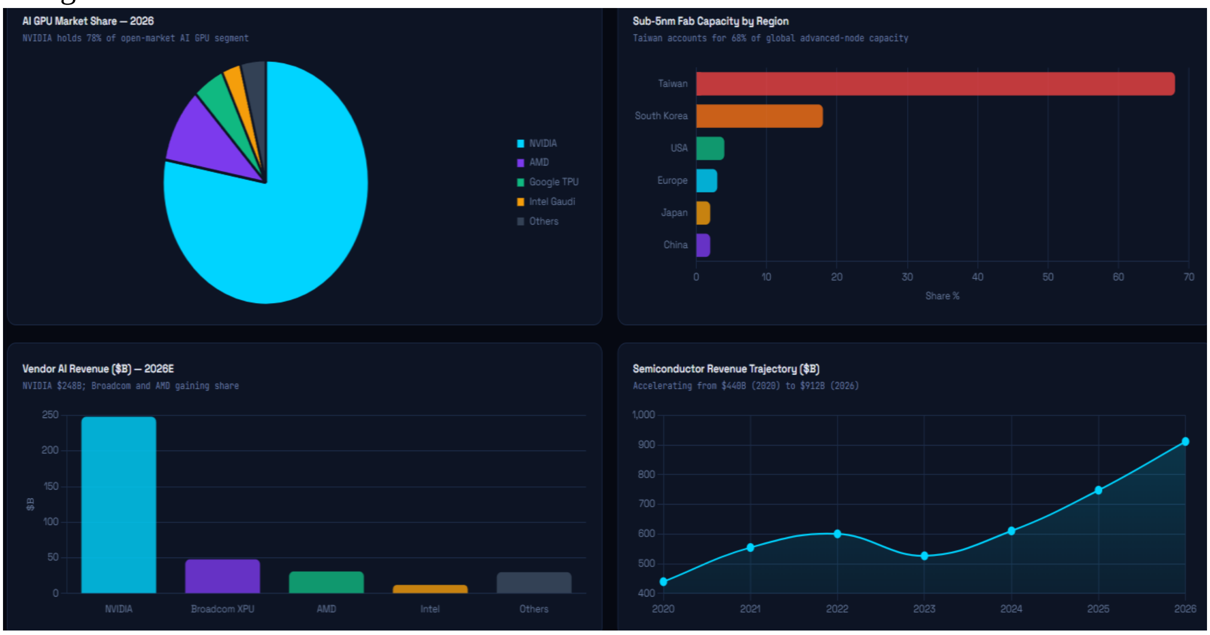

Figure 2. Global AI Chip Market Concentration and Advanced-Node Semiconductor Capacity.

The figure 3 shows AI data centers' 2026 energy mix and the geographically concentrated but rapidly rising power demand and nuclear-backed capacity growth projected through 2030.

Figure 3. AI Data-Center Energy Mix and Power Demand Dynamics, 2022–2030.

The figure 4 summarizes 2026 geopolitical supply-chain vulnerabilities across critical compute nodes highlighting highest risk at Taiwan advanced nodes and advanced packaging and shows the United States leading in export-control restrictions.

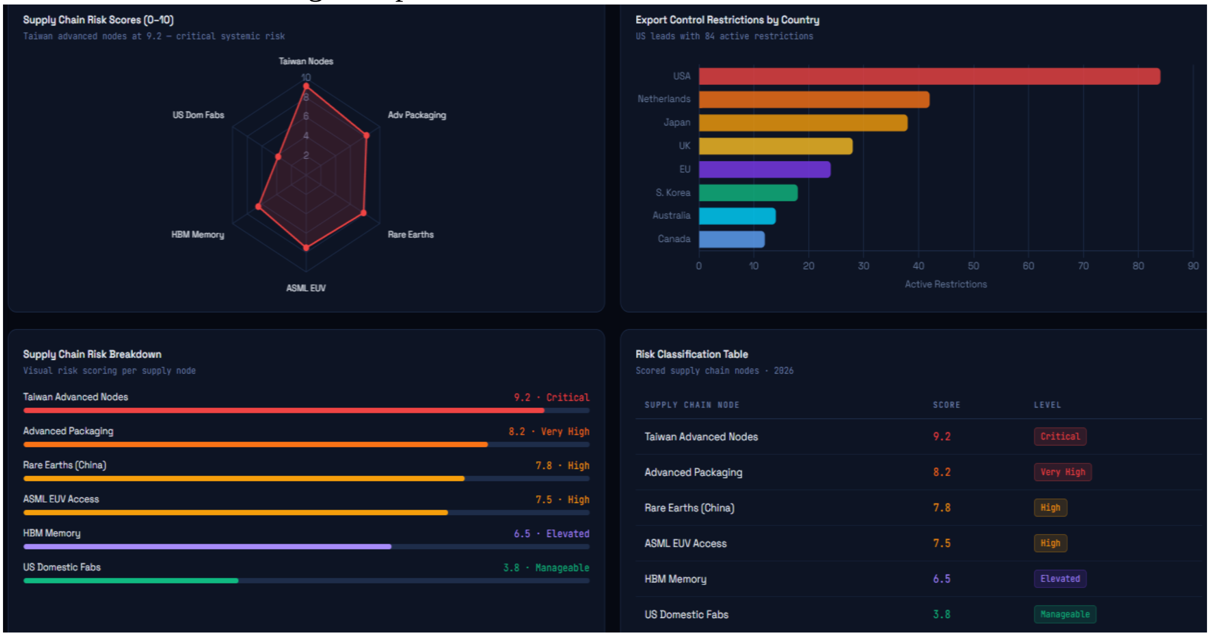

Figure 4. Geopolitical Supply-Chain Risk and Export-Control Intensity Across Critical Compute Nodes, 2026.

The figure 5 shows how AI is boosting GDP across countries, raising sector-level productivity, and driving rapidly increasing AI investment in emerging markets from 2022 to 2026.

Figure 5. AI-Driven Global Growth: Country-Level GDP Boost, Sector Productivity Gains, and Emerging-Market AI Investment, 2022–2026.

The figure 6 presents bull, base, and bear scenarios for 2026–2030 trajectories of semiconductor revenue, AI compute demand, data-center power use, and AI GDP contribution.

Figure 6. Scenario-Based Forecasts for Semiconductor Revenue, Compute Demand, Power Use, and AI GDP Contribution, 2026–2030.

5. Discussion: Implications and Limitations

The results from this analysis suggest that computing is now a key economic resource, on par with energy and transportation system, which is due to the fact that frontier AI outcomes rely on uninterrupted access to semiconductors, energy, money, and implementation of the favourable policies [1][2][8][15]. Thus, growth of AI investments, semiconductor revenues, and electricity consumed by data centers is considered as a sign of a wider production transformation, competitive capability, or strategy of the government in general, but not in respect to the AI industry in particular [5][9][22]. The first conclusion is related to the economics policy: the countries that manage to integrate IT infrastructure, develop energy systems, get access to semiconductors, and work out their industrial policy will benefit from higher level of production while others would remain dependent on imported computing even if they have significant software potential [21]. The GDP dynamics in the USA, India, China and UAE indicates that coherence of digital ecosystems plays an increasingly important role in gaining competitive advantage, which makes computing policy one of the most important aspects of development policy [19][20][22].

Another important point about industrial structures is that NVIDIA dominates in the open GPU market, along with TSMC's pivotal role in advanced nodes fabrication and being able to develop custom chips, as now they use only a small number of partner manufacturers and designers [8][9][12]. On the one hand, this concentration creates better conditions for scale and learning conditions. However, it exposes smaller companies and countries with slower development speed to threats of supply interruptions, costs increase, and lock-in effect [10][13]. Energy results tell us that increased AI operations make the relationship of the AI market growth with the grid changes and firms' low-carbon capacity stronger, bringing additional challenges on the path of climate transition in the systems where variable renewables do not guarantee reliability [15][16][17]. From the geopolitical point of view, the dependency on Taiwanese production, export control choke points, and raw materials turns the computing power directly into a part of strategic rivalry.

Nonetheless, the study has limitations. It employs a range of secondary and contracted data, a portion of which is model based and hence creates problems with comparability and estimation uncertainties. Results from scenarios will be sensitive to assumptions about energy build-up, regulation, semiconductor offerings, and geopolitical stability, it is important to treat the study as more directional than deterministic. Lastly, the framework tends to neglect coverage on labour markets, inequality, and non-electricity environmental impacts, suggesting that more work is needed to integrate social and institutional aspects.

6. Conclusion and Future Work

The researchers assert that compute economics is an overarching concept to understand the global political economy of AI by 2026 as, nowadays, AI development relies on semiconductor capacity, energy supply, and geopolitical control, not just on the software innovativeness [1][7][9][14]. Proofs of sustained AI capital investments, increasing energy consumption of data centers, growing revenue from semiconductors, and measurable contribution of AI to global GDP confirm that now expansion is structural leading to the change of economic priorities and industrial capabilities [22]. However, the outcome of this process is rather unequal: a few entities that have access to investment, subsystems of modern chips, well-equipped power grids, and industries that positively affect economic development will benefit, while others that suffer from restricted resources will undoubtedly lag behind [10][20][21]. Hence, the outcome of compute economics will generally depend on the decisions made by businesses and governments regarding the supply chain concentration, burden on the energy system, and privatization of export controls [11][13].

The paper illustrates that analysis restricted to a specific sector has become insufficient in the context of evaluating the impact of artificial intelligence on the economy. Integration of macro-level variables such as macroeconomic indicators, sectoral structure of the semiconductor market, energy system metrics, and geopolitical risk into a unified system, and its implementation in the Compute Economics 2026 Dashboard, allows providing a detailed description of the AI economy and a methodological platform for its measurement [8][9][15][19].

Longitudinal research is necessary to see if the prediction about compute demand, revenues from the semiconductor industry, and GDP gains resulting from the AI sector are justified in practice [3][8][22]. Social and institutional modules that cover labor market adjustment, inequality, and state capability need to be included in the project so that one can evaluate the consequences of building up the infrastructure [21]. Finally, environmental modules must go beyond electricity needs to include water usage, carbon intensity, land use impact, and hardware lifecycle waste [16][17].

On a larger scale, the emergence of compute economics signals a change in the way technology advances are perceived: one can no longer distinguish between the technology of artificial intelligence, or AI, and the physical and geopolitical systems responsible for the production and usage of compute technology, and that economic supremacy will increasingly depend on who is capable of acquiring, managing, and developing this key resource in an efficient manner.

References

AI Supremacy. (2025). *Burn baby burn: Exponential demand for compute supercycle*. AI Supremacy.

AI World Journal. (2026). *The AI productivity supercycle: Infrastructure, agentic software, and the great market divergence*. AI World Journal.

Association for the Advancement of Artificial Intelligence. (2024). *Compute North vs. Compute South: The uneven possibilities of compute-based AI governance*. AAAI.

Chow, P. (2026). *AI's next super-cycle: Financing the compute, power and infrastructure behind 2026*. LinkedIn Pulse.

EBC Financial. (2026). *AI infrastructure and energy supercycle: Market outlook 2026*. EBC Financial Group.

Elsevier. (2025). *Sustainable AI infrastructure: A scenario-based forecast of water footprint under uncertainty*. Elsevier.

EXANTE. (2025). *AI CapEx supercycle and energy infrastructure*. EXANTE.

Expert Network Calls. (2026). *Semiconductor market outlook: Key trends and challenges in 2026*. Expert Network Calls.

Gannatti, C. (2025). *The compute capital supercycle: AI's silent infrastructure revolution*. WisdomTree.

IEEE. (2025). *AI-RAN: Transforming RAN with AI-driven computing infrastructure*. Institute of Electrical and Electronics Engineers.

International Data Corporation. (2026). *Quarterly global AI compute and infrastructure forecast*. IDC.

International Energy Agency. (2026). *Data center energy demand and global power system impact report*. International Energy Agency.

Kania, E. B. (2025). *Technological competition and the semiconductor supply chain: A geopolitical assessment*. Center for a New American Security.

Pan, S., & Guo, K. (2025). *AI-driven occupational structural transformation and service-oriented manufacturing*. Jinan University & Sun Yat-sen University.

Perceptive IC. (2026). *Geopolitics reshaping the global semiconductor supply chain architecture*. Perceptive IC.

Perceptive IC. (2026). *The five-year AI super-cycle and the return of industrial strategy*. Perceptive IC.

Sanchez, V. (2025). *GeoCoded special report: State of global AI compute (2025 edition)*. GeoCoded.

Semantic Scholar. (2026). *Concentrated siting of AI data centers drives regional power-system stress under rising global compute demand*. Semantic Scholar.

Semiconductor Industry Association. (2026). *Global semiconductor revenue and fab capacity report*. Semiconductor Industry Association.

SJAIBT. (2024). *Data sovereignty in global AI-blockchain infrastructure*. *South Journal of Artificial Intelligence and Blockchain Technology*, 2(1), 45–61.

Sourceability. (2026). *Geopolitics are reshaping semiconductor supply chain risk in 2026*. Sourceability.

SSRN. (2026). *AI success beyond capability: Governance, energy, and trust as strategic determinants*. Social Science Research Network.

Stanford University. (2025). *Artificial intelligence and economic growth: A neoclassical framework*. Stanford University.

Toward a global regime for compute governance: Building the pause button. (2025). arXiv.
Z2Data. (2025). *22 critical supply chain risks to watch for in 2026*. Z2Data.