

Europe's AI Compute Gap: Why Strategic Leverage Matters More Than Hyperscale Parity

The Economy Research Editorial^{1,2}

¹The Economy Research, 71 Lower Baggot Street, Dublin 2, Co. Dublin, D02 P593, Ireland

²Swiss Institute of Artificial Intelligence, Chaltenbodenstrasse 26, 8834 Schindellegi, Schwyz, Switzerland

Abstract

This paper challenges the dominant narrative of Europe's strategic weakness in AI: its inability to compete with American and Chinese frontier-scale compute in the field. It argues that the AI compute gap is real, growing and politically important, but its strategic consequences are frequently misdiagnosed. Europe is not excluded from access to compute capacity alone; it is exposed to dominance in a narrow subset of domains, namely, reliance on concentrated foreign cloud platforms, inability to capture downstream value streams and erosion of upstream technological strength where it otherwise is a relatively privileged actor. A strategy that merely strives for numerical parity of brute force compute capacity would impose a burden both on Europe's fiscal capacity and its energy consumption, both of which are predisposed to disadvantages relative to its key rivals. Instead, it should strengthen its remaining irreplaceable upstream positions within the value chain, including ASML, IMEC, ARM, EuroHPC, AI Factories, and secure and sovereign compute for the public-sector, industrial and defense workloads. It turns then to defense as a more immediate strategic concern than the development of a rival civil frontier compute platform, and to the strategic imperative of asymmetry - reliance where ownership is unnecessary, sovereignty where dependence is not feasible, and preservation of upstream bottlenecks where necessary to make the EU indispensable.

1 Introduction - The Wrong Lesson from Europe's AI Compute Gap

The dominant framing of Europe's predicament in the AI debate is that because leading-edge AI is increasingly compute-hungry, Europe needs to close its compute gap with the US and China. A recent Bruegel intervention gives one of the clearest articulations of this framing: Europe needs to see AI computing as a strategic deficit to address with a catch-up approach.^[1] This framing, however, misses the structure of the industry that it claims to analyze. It assumes strategic power is synonymous with having the largest downstream compute clusters and that reliance arises simply from an absence of domestic training capacity. This premise is faulty. In today's AI economy, access, contractual control, upstream bottlenecks, power costs and alliance considerations matter just as much as gross installed compute. Europe's problem is therefore not primarily the quantity of compute it owns, but that policy discourse has all too frequently equated numerical inferiority at one layer of the tech stack with system-level strategic insignificance.

This distinction is, if anything, even more crucial between 2023 and 2026. Frontier AI has become more corporate, more capital-intensive and more geographically concentrated. Stanford's 2025 AI Index reported that nearly 90% of notable AI models in 2024 came from industry, not academia, that US-based institutions released 40 models compared with China's 15 and Europe's three and that the compute required to train models was doubling about every five months.^[2] The same structural pattern is visible in corporate capital spending: Microsoft announced roughly \$80 billion for AI-enabled data centres in FY2025; Alphabet guided capital expenditure of about \$75 billion for 2025 before later upward revisions, Meta projected \$60-65 billion for 2025 and Amazon's capital expenditure reached \$131 billion in 2025 while 2026 expectations moved toward approximately \$200 billion.^[3] Relative to these numbers, Europe's principal public policy response (the Commission's 20 billion gigafactory fund in InvestAI) is substantial by EU policy standards, but it is nowhere near volume-equivalent to the hyperscaler race and it therefore needs to be assessed for strategic leverage, not narrative parity.

This reframing reveals a second mistake: The idea that one can have absolute sovereign control of the entire semiconductor and AI value chain is erroneous. The OECD's 2025 mapping of the semiconductor value chain confirms that it is inherently globally distributed and interdependent, with particular inputs becoming concentrated in specific geographies.^[4] Even the US, which appears closest to having a complete AI stack, relies upon foreign nodes that are critical: Dutch lithography, Taiwanese advanced fabrication and Korean high-bandwidth memory. Reuters described ASML as having a monopoly on extreme ultraviolet lithography systems, TSMC as the world's main producer of advanced AI chips and SK Hynix as having the largest market share of HBM chips in 2025.^[5] What these developments demonstrate is not that Europe should necessarily attempt to replicate US leadership at a later stage, but rather that autarky is futile and choke points remain critical.

China and Japan make the point from the opposite end. China has built up domestic capabilities to counteract foreign dependence with a whole-of-state effort, including the third state-funded semiconductor fund launched in 2024, which has registered capital of 344 billion yuan (\$47.5 billion).^[6] Japan's recent push into advanced logic already entails state funding for Rapidus of over 2.3 trillion yen, but mass production is still tentatively forecast for FY2027.^[7] Neither China nor Japan offers an example of easy state-led catch-up; both

show that reconstructing a domestic capability when one is outside the leading edge is expensive and risky. Europe should be learning from this caution, not embracing a romanticized version of it. A late entry into a compute race, coupled with higher power prices, fragmented capital markets and numerous competing public expenditure priorities, suggests that an attempt to mimic brute-force catch-up will be the most expensive possible formulation of strategic autonomy.

The foundation for a more sensible European strategy lies in Europe's existing strengths. ASML continues to play an indispensable role in leading-edge fabrication. IMEC in Belgium has acquired even greater importance as a cooperative research platform, with its NanoIC pilot line, under the Chips Act, having pooled €2.5 billion and housing one of the few globally available High-NA EUV tools for sub-2nm development.^[8] If the remit is widened from the EU to Europe as a whole, ARM also has significance as a crucial architectural designer, whose influence on data-centre CPUs is rapidly expanding. In sum, Europe has a role in the AI value chain; it is a specialized role. The wisest approach is not an attempt to establish immediate parity in owned frontier compute, but to secure and strengthen European capacities which cannot easily be substituted elsewhere while still ensuring access to allied cloud and AI services where direct replication is uneconomical. This is not a passive strategy, but rather asymmetric statecraft.

At its core, the argument is quite simple. Europe does not need to win a computing race in the same way Washington or Beijing do to be a relevant actor in AI. It only needs to ensure compute can be acquired on acceptable terms, sensitive workloads can run in trusted sovereign environments when necessary and upstream European choke points are preserved and upgraded rather than allowed to atrophy into legacy technologies. The remaining sections unpack this contention in three parts: what is the compute gap, would increased domestic compute have benefits justifying its costs in Europe's specific energy and fiscal circumstances and whether, amid the security pressures of 2025–2026, defense readiness has now taken precedence over the symbolic search for AI hegemony.

2 Computing Power Gap: Ownership, Access, and Strategic Leverage

The compute gap is real and an honest discussion of the situation must begin by acknowledging it clearly. As estimated by the IEA, 415 TWh was consumed globally by data centres in 2024: 45% in the United States, 25% in China, and 15% in Europe.^[9] The IEA predicts this demand will more than double to about 945 TWh by 2030, with most of the growth occurring in the US and China, combined at around 80%.^[10] On a per-capita basis, the imbalance is even more severe: 540 kWh of per-capita US data-centre demand, about an order of magnitude higher than that of any other region. These are not negligible disparities; they indicate entirely different positions in the global architecture of AI computing. The frontier also mirrors this disparity; according to the AI Index, in 2024 US institutions released 40 seminal AI models, China 15, and Europe just three.^[11] Both the electrical load and the output that compute generates exhibit significant gaps.

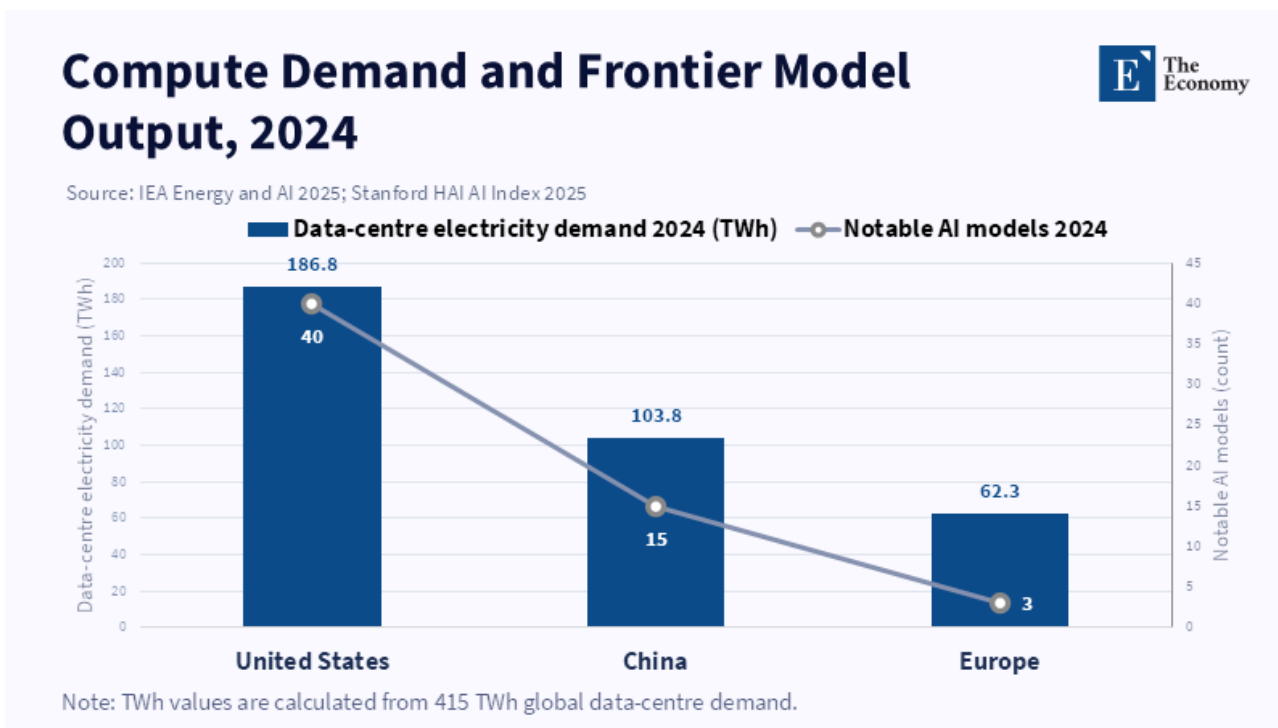


Figure 1: Compute scale and frontier-model output are concentrated in the same geopolitical centres, but this does not automatically mean Europe lacks usable access.

However, a gap in owned computing capacity does not automatically translate into a gap in usable compute access, even though it can still weaken bargaining power, value capture and resilience in strategically sensitive sectors. At this point, much of the European argument devolves from observation into exaggeration. As the OECD 2025 review on cloud market competition shows, cloud services generally do not operate at a national level, but rather at a regional one;^[12] it is an infrastructural, not ideological, choice: servers in the same region can deliver data quickly enough for most workloads, and large users want infrastructure distributed across multiple states to meet many demands, not restricted to a single nation. The article also indicates that cloud infrastructure is technically available everywhere via the internet, provided certain sectors have adequate data sovereignty, a fact that cannot be overlooked. Europe is not shut out of compute because it lacks indigenous hyperscalers; it is integrated into a regional cloud market in which foreign-owned providers already operate substantial capacity on European soil.

There is empirical evidence for this claim. AWS offers a dense network of European regions, including Ireland, Frankfurt, London, Stockholm, Spain, Milan, and Zurich; Google Cloud explicitly notes that its 43 global regions will facilitate foundational and agentic AI applications while simultaneously allowing data residency customization; Microsoft is expanding its European operations further, with over 80 global data centre regions including a steady rollout of new and improved facilities in numerous European nations.^[13] As such, for most companies the primary barrier to AI progress is not how to get a model trained by a European hyperscaler; it is finding affordable compute with adequate performance, contracts and compliance rules in the right region. An economy can lack top-tier ownership and still have massive operational access to computing power; Europe finds itself in such a situation.

Nor is it clear that Europe consistently faces structurally punitive prices.^[13] According to its own list, on

April 4th the effective price for a p5.48xlarge with eight H100s was \$34.608 in Northern Virginia, but \$31.464 in London and Stockholm. The very same table offers P5E H200 capacity blocks for a constant price across multiple non-US regions, including London and Stockholm, while the price on the table for the P5EN capacity block was lower in European regions than in the US. While these figures do not resolve questions of queue times, negotiated discounts, egress fees, legal risk or capacity scarcity, they do show that Europe does not uniformly face a simple nationality-based compute surcharge; the deeper issue is strategic dependence on foreign platforms rather than straightforward market exclusion.

This differentiation highlights a more controlled vulnerability. Europe’s vulnerability derives from foreign political will being able to restrict access, the high switching costs of locking into foreign platforms and the inability of sensitive sectors to endure extra-territorial supervision. The OECD shows these concentrations persist.^[15] Their report says Amazon and Microsoft together have a 70% to 80% market share in some of Europe’s largest countries, with AWS holding 46% in France compared to Azure’s 17% and a 45% to 35% split in the Netherlands, while in the United Kingdom the figure reaches 80%.^[16] Their table lists the problem of customer lock-in and high barriers to entry as structural characteristics of the market. This means Europe’s vulnerability has less to do with a raw shortage of compute capacity than with a governance problem created by market concentration and lock-in. While serious, this doesn’t mean the only answer is to develop a European bloc of hyperscalers; it requires attention to interoperability regulations, public procurement rules, sovereign clouds and bargaining power stemming from upstream strengths.

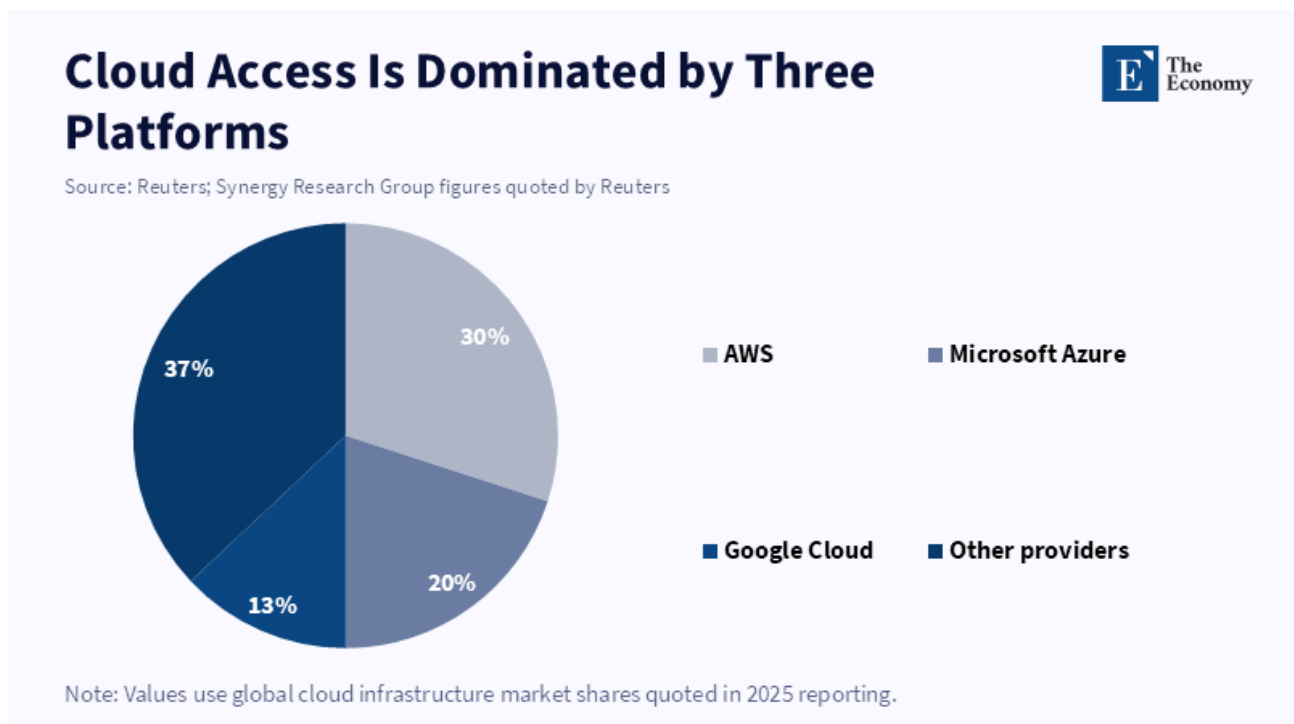


Figure 2: The strategic risk is not simple exclusion from cloud services, but dependence on a narrow platform layer.

It is these upstream strengths that form the most critical aspect of the equation; they are too often relegated to a footnote in the pursuit of glamorous model training. The OECD’s mapping of the semiconductor industry underscores how the entire value chain is interrelated and how various countries specialize rather than vertically

integrate.^[17] It shows how EU companies play a vital role in virtually every part of the value chain except chip fabrication (based on previous work by the JRC); this residual importance is not trivial and provides Europe with its most significant source of bargaining power. ASML is the exclusive provider of EUV lithography systems, a fact that grants it exceptional geostrategic importance, as all advanced fabrication relies upon it. IMEC, through its NanoIC pilot line, extends this importance from hardware to process development, prototyping and ecosystem orchestration.

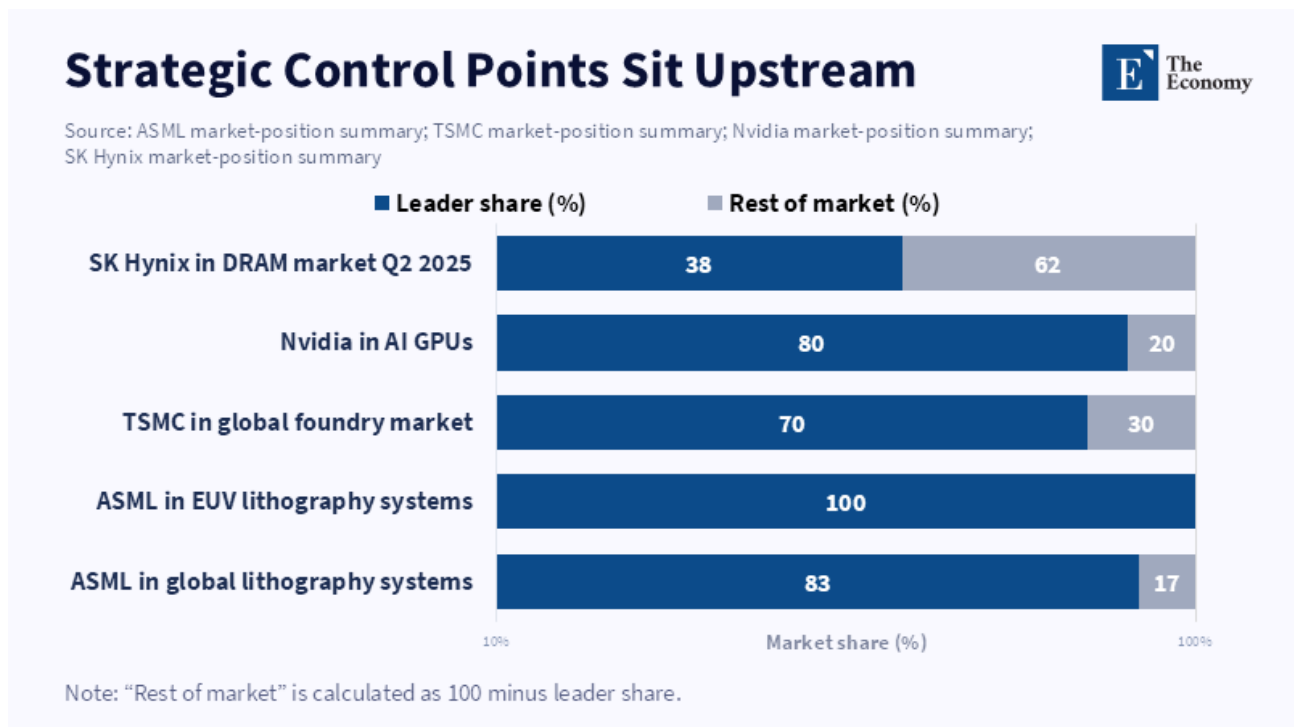


Figure 3: Europe’s leverage lies less in matching hyperscale compute than in preserving bottlenecks others cannot bypass.

The parallel with critical raw materials provides further clarity. Reuters published stories in 2025 and 2026 about China’s restrictions on exports of gallium, germanium and rare earth metals leading to spikes in prices and shortages; the EU has responded by forming joint stockpiles to offset dependence. The lesson here isn’t to embrace each concentration.^[18] It is to recognize that upstream chokepoints generate bargaining power and determine the allocation of rents and risk. Should Europe allow its critical positions in semiconductors to diminish while focusing on costly efforts to replicate the US-China compute race, it would sacrifice its strongest form of leverage for the weakest element in the value chain. This would hardly be strategic autonomy; instead, it would be a strategic undermining of Europe itself.

There is one other vital point: Europe already possesses a system of public infrastructure that is more fitting to its structure than the pursuit of its own hyperscalers. The AI Factories program initiated by the Commission under EuroHPC provides 19 AI Factories and 13 Antennas that are currently operational and will procure at least nine new AI-optimized supercomputers, tripling EuroHPC’s existing AI computing capacity.^[19] Access to this system is open to European industry, research, academia and public authorities and priority is given to startups and SMEs. This is not equivalent to Silicon Valley’s computing capacity but rather an interconnected public computing system designed to maximize diffusion and access. It is possible that this model offers a more

appropriate means to spread capacity across the continent and promote experimentation rather than trying to make Europe another Amazon or Microsoft to remain a player.

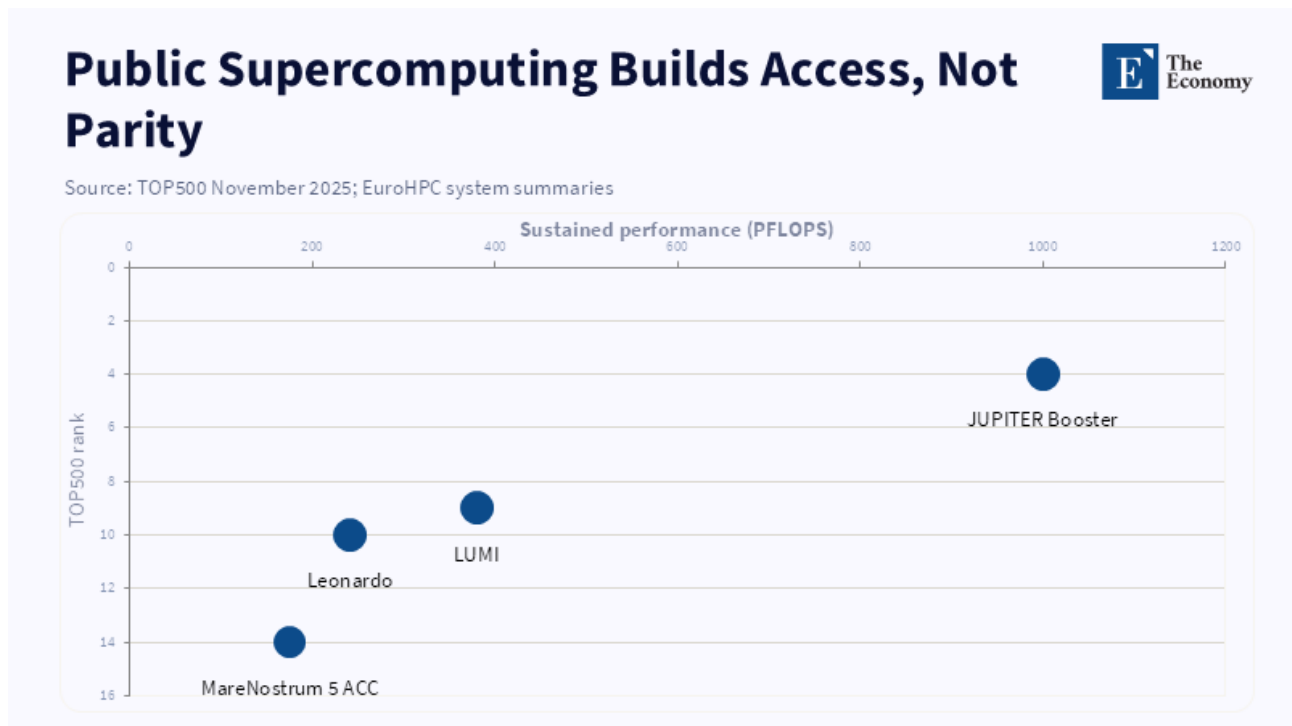


Figure 4: Europe’s leading public supercomputers provide serious shared capacity, but their role is access and experimentation rather than private hyperscaler-scale AI ownership.

What this implies is not that Europe can afford to live with the compute gap but should do so selectively rather than imitatively. The strategic objective is to ensure secure access, dilute concentrated dependency and protect Europe’s upstream leverage rather than to pursue symbolic parity in hyperscale capacity at any price.

3 Cost vs. Benefit of Additional Computing Power Under Europe’s Energy Constraints

Therefore, the core policy question is not whether more computing has value. It clearly does. The question is whether closing the gap via increased computing has marginal strategic and economic benefits that are worth the marginal cost to Europe’s specific context. Commission documents show that Brussels is indeed signaling capacity acceleration; InvestAI aimed to mobilize 200 billion euros in total, including 20 billion euros for AI gigafactories, while the forthcoming Cloud and AI Development Act is intended to at least triple the EU’s datacentre capacity within 5–7 years and to cover the needs of businesses and public administrations fully by 2035.^[20] The AI Factories page, moreover, notes that gigafactories are intended to have ~100,000 advanced AI chips each and to demand massive amounts of power. This is understandable as a policy aspiration but it requires much stricter scrutiny as an economic strategy. Europe is not deciding to add some compute capacity. Europe is deciding to socialize a hugely expensive arms race in a market in which other actors are already spending by an order of magnitude more.

Energy remains the first and least avoidable constraint. In H2 2025, average EU household electricity

prices were 0.2896/kWh, near 2024 highs and well above pre-crisis levels; for non-household consumers, it was 0.1837/kWh.^[21] Eurostat shows that household electricity prices rose by 36% on average across the EU between 2019 and 2024; in the US, they rose by 26%.^[22] Similarly, a recent IEA report found that 2023 wholesale electricity prices in Europe, while retreating from 2022 peaks, still stood double pre-crisis levels, while prices were only 15% higher in the US. These figures are not proof that power cannot be made available to the continent but they indicate Europe is approaching a vast scaling of AI infrastructure from a significantly weaker starting point regarding energy costs than its main rival. Furthermore, powering added computing through expensive electricity has the risk of appearing and proving to be more costly and less justifiable in a politically sensitive energy environment where households and businesses are already facing elevated bills.

The second constraint is not merely the general price level across Europe, but the localized intensity of the load. The European Parliament’s 2025 briefing to the Committee on Industry, Research and Energy stated that datacentres represent around 3% of overall EU electricity consumption but over 20% in Ireland,^[23] which, according to Ireland’s Central Statistics Office, saw data centre usage represent 22% of Ireland’s total metered electricity consumption in 2024, up from 5% in 2015.^[24] In a separate report, the IEA noted that in key EU hubs like Dublin and Amsterdam, further expansion has been postponed due to a lack of grid capacity and the difficulty of integrating new high-power demand sites.^[25] It is evident then that both electricity price and grid constraints are fundamental to the economics of European AI datacentres, alongside siting, planning and politics issues that complicate integrating energy-consuming facilities into highly strained grid nodes. These are not theoretical problems; they already affect Europe’s most established digital locations.

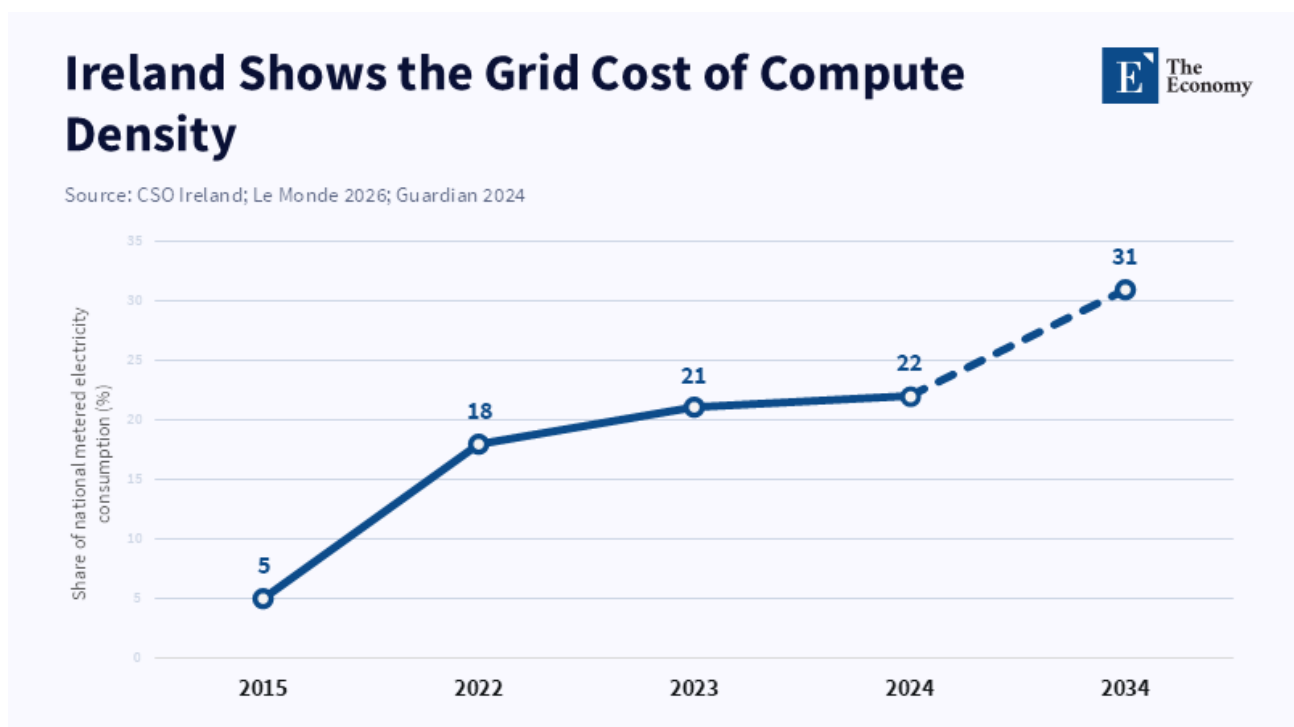


Figure 5: Ireland shows how data-centre expansion becomes a grid-management problem before it becomes a sovereignty success story.

The context for this point is the particularly high capital intensity of training frontier models-Microsoft has pledged approximately \$80 billion in FY2025 for AI data centre buildout; Alphabet, \$75B for 2025 (which

has already been upward revised); Meta, \$60-65 billion; Amazon’s 2025 and 2026 capital expenditure has risen sharply in each year to the \$200B mark for the latter.^[26] Even factoring in that this expenditure does not translate one-for-one into AI model training costs, the capital sum is simply an order of magnitude higher than the EU’s €20 billion gigafactory fund and indicates the realities of hyperscale competition, which demands immense, sustained capital deployment over years rather than a well-meaning public announcement.

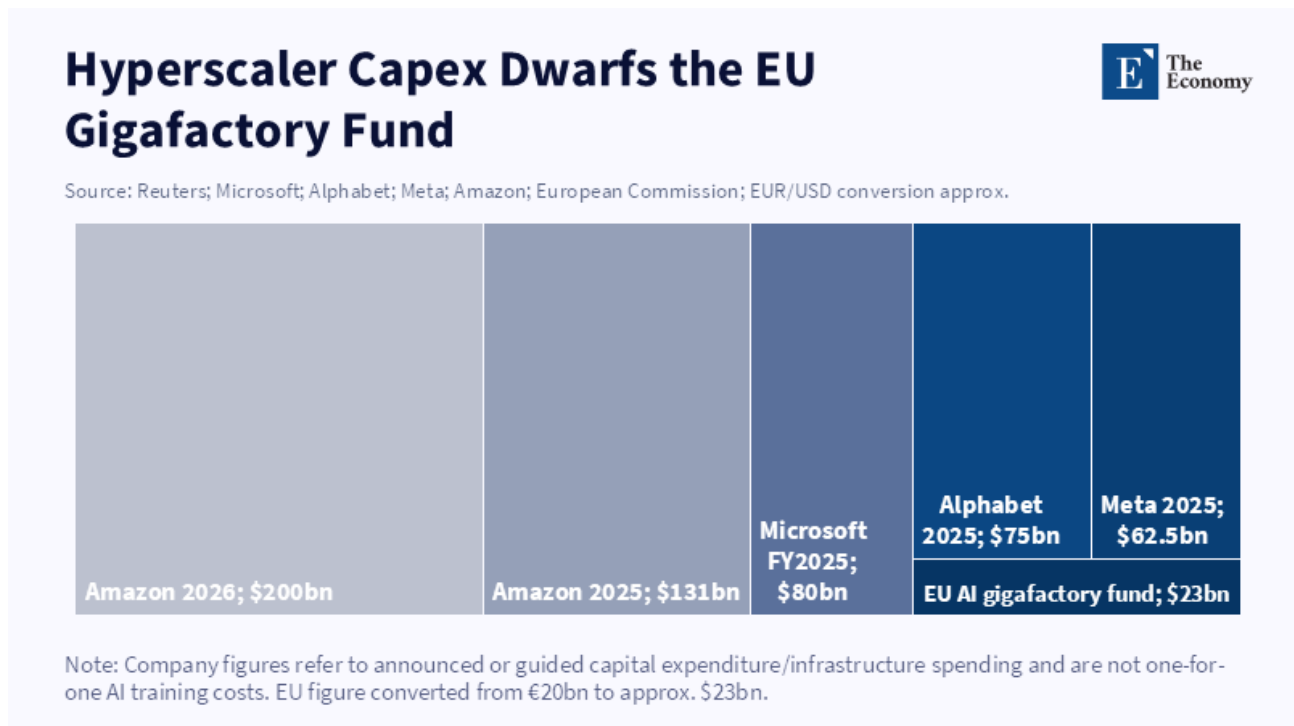


Figure 6: The financial benchmark for frontier AI infrastructure is being set by hyperscaler balance sheets, not public industrial-policy funds.

Of course, the advantages of more computing capacity aren’t limited to the high end of frontier AI training. Inference, industrial AI applications, scientific research, the formation of startups and even leveraging with foreign cloud providers all benefit from available computation. This argument is entirely valid, but the leap to suggesting a need for parity is faulty. Europe’s own proposals, from the Commission’s Gigafactory initiative’s plans for open access and the tripartite increase to the EuroHPC AI capacity, to EU sovereign AI efforts in sectors such as Mistral’s recent funding round to purchase 13,800 Nvidia chips for a data centre near Paris, are selective and modular, rather than direct, and proportional imitations of US hyperscalers.^[27] The fact that Europe is focusing on specific segments such as sovereign compute or particular aspects of the AI infrastructure stack, rather than a blanket increase in raw capacity, is a strength rather than a weakness.

Furthermore, a realistic analysis should also address the likely distribution of benefits from increased infrastructure development. In practice, the principal gains from expanded AI infrastructure are most easily captured by large incumbent firms that already possess the software, procurement, deployment and commercialization capabilities Europe still lacks at scale. Stanford’s AI Index shows that industry accounted for 90.2% of notable AI models in 2024 and that Europe produced only three of them; under those conditions, indiscriminate infrastructure expansion risks leaving Europe with a larger share of the energy and capital burden than of the downstream model rents.^[28] The problem isn’t the provision of raw compute; it is the absence of complemen-

tary software, infrastructure and the economic environment that makes profitable model development feasible. Thus, any proposed addition of compute needs a well-defined workload-specific benefit case, rather than an undifferentiated national benefit.

The best counterargument is that at every phase of industrial transformation, it appeared that the necessary investment was unfeasible until network effects kicked in, and without scale now, Europe will always be locked out of the technology. While the potential for path dependence on technological inferiority is a real one, history has also shown that the worst aspect of industrial policy is backing the wrong part of the technological stack. Europe needs to invest, but it needs to invest wisely. The greatest returns for Europe in AI and semiconductors would appear to lie in consolidating lithography excellence, scaling pilot lines and shared facilities, expanding packaging and compound semiconductors, developing power electronics, fostering edge-cloud convergence and providing sector-specific, sovereign computing solutions where there is a clearly definable user base with the potential to generate returns in the short to medium term. This allows for shaping the standards, commanding the rents and addressing acute dependencies. Massive, open-ended increases in data centre capacity in an energy-constrained continent have demonstrably diminishing returns.

Europe needs a disciplined increase, rather than inaction. More compute should be added in a tiered, geographically sensible manner: public shared compute to support research and startups; sovereign, secure compute to ensure state and defense capabilities; and a market-led, power-abundant, cost-efficient expansion for commercial services. This gradualist approach should recognize, as the EU Commission's own cloud policy already begins to, the link between the edge and the cloud, and integrate the provisioning of computing capacity with broader energy systems and sustainability considerations. The temptation to transform a prudent access-enabling policy into a prestige continental project that mimics expensive global efforts while disproportionately increasing costs should be resisted.

4 Why Defense Readiness Is More Urgent Than Civilian Frontier-Compute Parity

If the cost dimension of the compute race looks forbidding enough, the strategic opportunity cost looks almost impossible to avoid, once defense is pulled into the frame. The EU's security environment of 2025–26 intensifies the need for rearmament, readiness and autonomous defense-industrial capability. The Commission's defense page makes plain that the ReArm Europe plan and Readiness 2030 framework are geared towards mobilizing 800 billion; including 650 billion that may be raised through fiscal space derived from a 1.5 percent increase in annual EU GDP spending on defense over four years and 150 billion via a SAFE loan facility.^[29] The Kiel Institute and Bruegel estimate in early 2025 that defending Europe would cost over 250 billion per year extra for defense spending, and an additional 300,000 soldiers, without the United States.^[30] These are not minor adjustments financed via administrative creep; these figures represent massive resource reallocation – within fiscal, industrial and political constraints.

The realized figures in terms of spending already signal this. According to SIPRI's reports, Europe's military

spending rose 14% in 2025 to 864 billion dollars, with European NATO members-all 29 of them-spending 559 billion.^[31] The report details how European NATO military spending rose faster than any time since 1953, both due to war and to European aspirations to be self-sufficient militarily. Likewise, the Council of the EU states that the total spending of EU member states in defense rose to 343 billion euros in 2024, a figure that is set to rise to 381 billion euros in 2025.^[32]

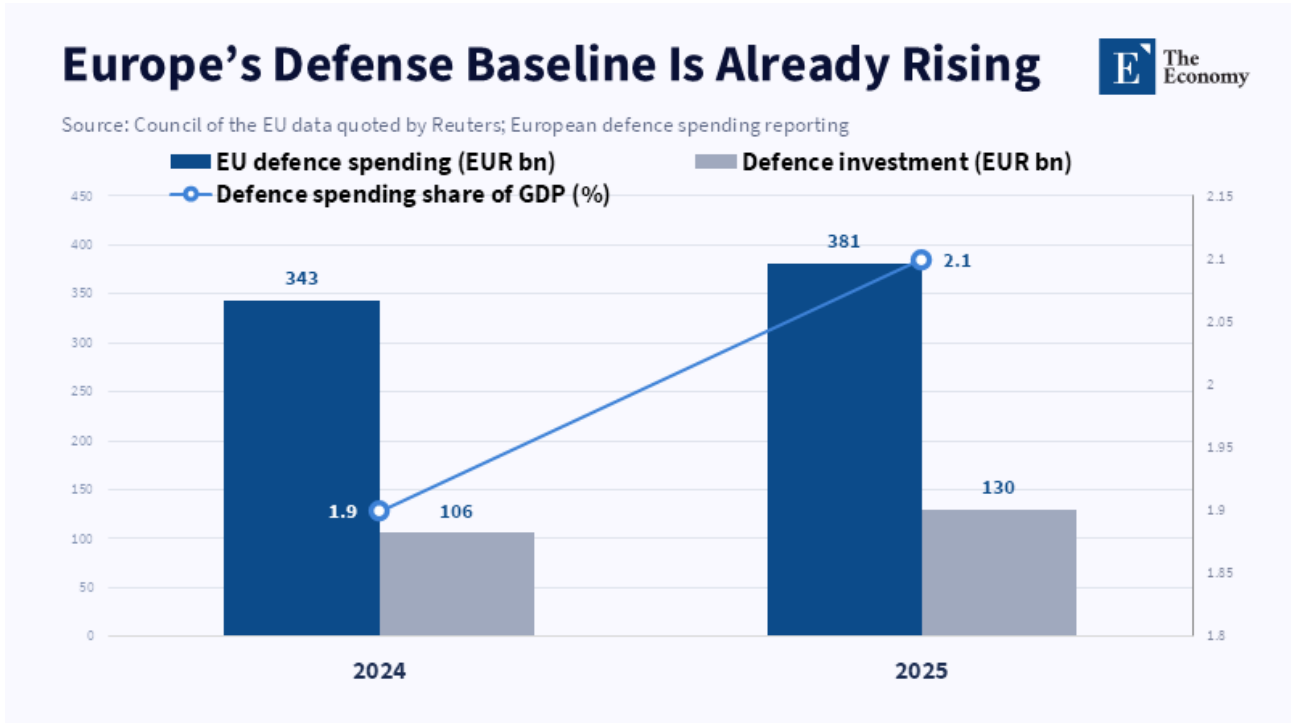


Figure 7: Europe’s fiscal room is already being pulled toward defence before any full civilian compute catch-up is financed.

Regardless of one’s view on the precise trajectory of the burden-sharing arrangements with the US, there can be no doubt that Europe now faces a security financing challenge of the first magnitude. Under these conditions, any demand that the EU spend large public sums to catch up in frontier computing capabilities requires a powerful justification. That justification is not currently present.

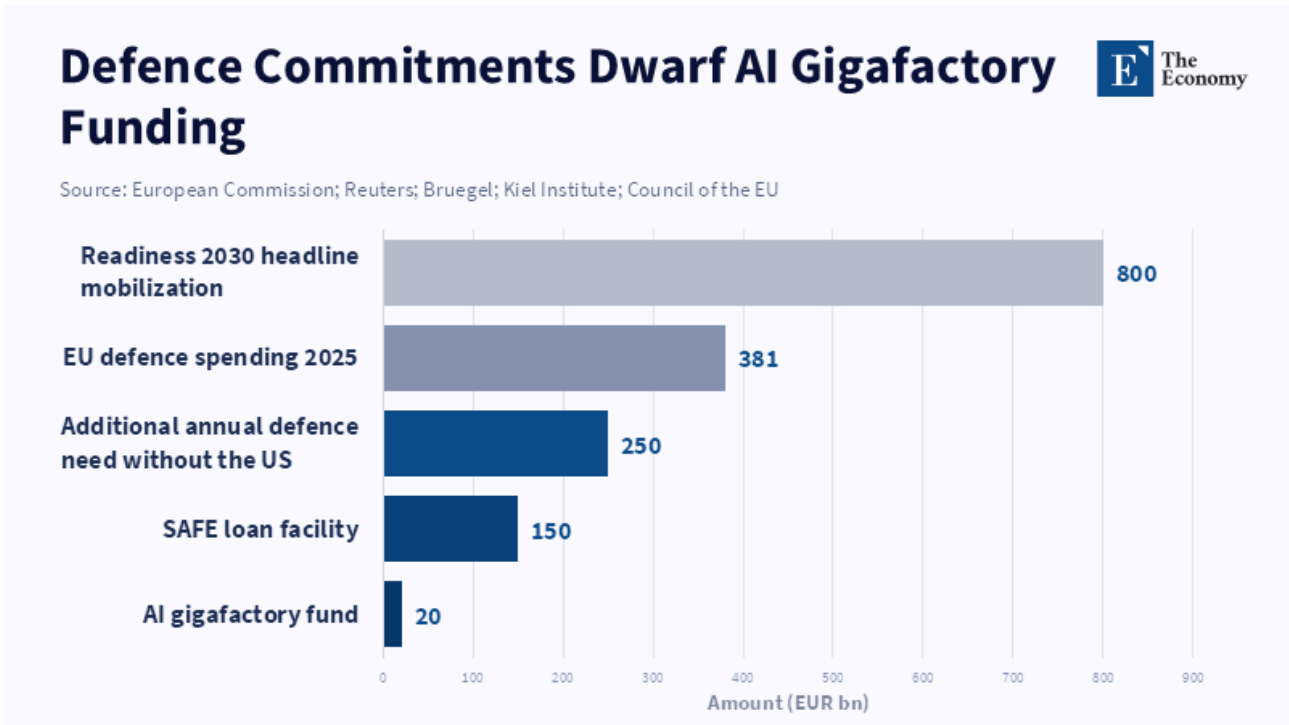


Figure 8: When defence needs are placed beside AI compute funding, the opportunity cost of a parity strategy becomes visible.

That does not mean that AI will not be of use to defense. Indeed, defense is the area where its capabilities can be readily found and it is a stated objective of the Commission’s defense readiness roadmap to establish capability coalitions in areas like AI, cyber and electronic warfare. DIANA, NATO’s innovation fund, was set up specifically for fostering dual-use innovations for the defense alliance.^[33] However, defense does not need to take the same form as a commercial frontier model arms race. Many tactically and strategically relevant defense and intelligence tasks depend on reliable inference, sensor fusion, logistics optimization, automation, target allocation, electronic warfare modification, the fine-tuning of specialized domain models on classified data. These functions require secure compute environments and cloud infrastructures but need not require Europe to create its own general-purpose foundation models and train them from scratch on European-owned hyperscale computing centres. Sovereign defense computing, while substantial in size, does not approach the massive demands made by commercial general AI parity.

An obvious counter-argument arises immediately: if Europe uses allied or foreign-owned compute in AI and defense, does this not deepen exactly the kind of dependency that strategic autonomy is supposed to obviate? For the most sensitive applications, the answer is absolutely yes. These workloads cannot be handed off without incurring legal, operational and intelligence risk. OECD’s study on competition in cloud computing already recognizes that government, health, and financial institutions will need to use domestic or specially managed cloud infrastructures in order to mitigate geopolitical risks and ensure security.^[34] This is an important point which disproves any simplistic claim about the irrelevance of sovereignty just because computing services are a global commodity; however, it does not imply that every task needs to acquire immediate sovereignty-rather, sovereignty needs to be applied where and only where it makes sense. Prudent strategies establish boundaries, rather than eliminate them.

These delineated boundaries are slowly becoming a reality in practice. As a report by Reuters in May 2026 demonstrates, Thales and Google Cloud struck a deal for a sovereign cloud in Germany run by Thales, which is 100% owned and controlled by this German company. France had earlier proceeded on similar terms: recognizing the technical asymmetry, while maintaining political and operational control in legal structures.^[35] This modular approach appears more viable for Europe than strict sovereignty everywhere—the commercial cloud market will continue to be dominated by foreign actors; however public, defense and national security functions require secure environments immune to extraterritorial oversight. In sum, defense computing does require sovereignty, but that requirement does not justify a generalized civilian race for frontier-compute parity.

From this perspective, the strategic choice appears clear. Europe needs to separate the three distinct layers that are frequently conflated in public debate: The first is the layer of highly sensitive sovereign workload in public administration, regulated sectors, intelligence and defense; these warrant the use of secure cloud, robust computing infrastructure and in certain cases dedicated on-site or ring-fenced hardware. The second is the layer of research, startup and industrial experiments; EuroHPC and AI Factories fit this layer, sharing access to computing power without pretending to general European hyperscaler parity. The third is the layer of scale-up services in the general market; here the role of allied and foreign providers will probably continue to be dominant and regulation, procurement and interchangeability of services may be more important than ownership per se. When these three layers are segregated, the push for a single compute gap closure objective loses its rationale. Europe has different needs and varied requirements for sovereignty. Its financial resources are not limitless.

The existence of European AI companies also suggests a strategy focusing on niche, rather than universal, ambitions. Europe is not devoid of capable AI companies. Mistral released its first reasoning model in 2025 and secured sufficient investment in 2026 to build a solid computing base, also for the French military.^[36] This is significant and proof that it is possible to build sovereign, dedicated capacity where there is demand and political will. However, it also shows the limitations of this catching-up paradigm: Mistral's computing power is modest compared to that of the big hyperscalers; it is appropriate precisely because it is specialized, controlled, and responsive to concrete and restricted European needs for computing infrastructure. From this, it is only possible to infer that Europe should focus on developing firms and technologies serving strategic niches rather than claiming that Europe should base its future on general computing hegemony. The rational strategy is focused: Europe needs to protect upstream supply chains, secure trusted access to foreign compute capabilities, fund dedicated sovereign compute infrastructure for needs that truly require it, and focus finite budget constraints on ensuring indispensability rather than expensive sameness.

The issue is ultimately political-economic. Europe's historical advantage has never been that of achieving hegemony by building an all-encompassing imperial technological enterprise. Its strength has been rooted in the creation of standards, institutional frameworks, specialized industrial capacities, and leveraging regulation within open but structured markets. This model has limitations, particularly today with rising great power rivalry. However, it offers a much better basis for an AI policy than simply emulating the US-China competition in an area where Europe is not in an analogous position. Defense now requires vast sums, energy prices remain politically sensitive, and capital markets remain fragmented. Europe's way back to relevance will therefore be a more subtle, less totalizing approach; the priority is securing strategic chokepoints, ensuring trusted access to

foreign cloud infrastructure, providing public, mission-oriented sovereign computing where sovereignty demands it, and using limited resources on capabilities in which Europe is uniquely placed, rather than expensive.

5 Conclusion - Indispensability, Not Compute Hegemony

The core error in much of the debate today is that it defines Europe's AI future as only one parameter: owned frontier compute. That parameter describes something real but not the entire geometry of strategic power. Europe's weakness in downstream AI compute is real, and will only grow as US hyperscalers and China's state champions continue their exponential scaling. But compute access is not ownership and dependence is uneven along the value chain. Europe still has clout in areas where others must rely on Europe's upstream function, particularly lithography, pre-competitive semiconductor research (IMEC), and the hardware-design ecosystem more broadly. These are important from economic and geopolitical perspectives.

It is for this reason that Europe should not embrace a strategy of prestige focused on brute-force parity of compute. This strategy would yield low marginal returns on a continent constrained by energy and public debt, and facing imminent pressure to rearm. The alternative to surrender or nostalgia is not more of the same, but pragmatic statecraft: doubling down on ASML and IMEC; leveraging EuroHPC and AI Factories to provide access; building secure sovereign compute for workloads that demand it; and leveraging allied, diverse clouds for the rest. Europe does not need to stake its future on an AI hegemony contest, it is poorly positioned to win. It does need to preserve and deepen the indispensable nodes that make others depend on Europe.

References

- [1] Bruegel (2025) *Europe Needs a Strategy to Close the Artificial Intelligence Compute Gap*. Bruegel.
- [2, 11, 28] Stanford Institute for Human-Centered Artificial Intelligence (2025) *AI Index Report 2025*. Stanford University.
- [3, 26] Microsoft (2025) *Microsoft to Invest \$80 Billion in AI-Enabled Data Centers in Fiscal 2025*. Microsoft; Alphabet (2025) *2025 Capital Expenditure Guidance*. Alphabet; Meta (2025) *2025 Capital Expenditure Guidance*. Meta; Financial Times (2026) *Amazon's AI Infrastructure Spending and Capital Expenditure*. Financial Times.
- [4, 17] OECD (2025) *The Semiconductor Value Chain: Recent Evolution and Policy Considerations*. OECD.
- [5] Reuters (2025) *ASML, TSMC and SK Hynix in the Global AI Chip Supply Chain*. Reuters.
- [6] Reuters (2024) *China Launches Third State-Backed Semiconductor Investment Fund*. Reuters.
- [7] Rapidus / Government of Japan (2025) *Rapidus Funding and Advanced Logic Production Timeline*. Rapidus / Government of Japan.
- [8] IMEC / European Commission (2025) *NanoIC Pilot Line and European Chips Act Support*. IMEC / Euro-

pean Commission.

- [9, 10, 25] International Energy Agency (2025) *Energy and AI*. IEA.
- [12, 15, 16, 34] OECD (2025) *How Europe Can Deploy Cloud to Strengthen Competitiveness*. OECD.
- [13] Amazon Web Services (2026) *AWS Global Infrastructure*; Google Cloud (2026) *Global Locations and Regions*; Microsoft (2025–2026) *Global Cloud and Data Center Regions*.
- [14] Amazon Web Services (2026) *EC2 Capacity Blocks for ML Pricing*. AWS.
- [18] Reuters (2025–2026) *China's Export Restrictions on Gallium, Germanium and Rare Earths*. Reuters.
- [19] European Commission / EuroHPC Joint Undertaking (2025–2026) *AI Factories Initiative*. European Commission / EuroHPC JU.
- [20] European Commission (2025–2026) *InvestAI and the Cloud and AI Development Act*. European Commission.
- [21] Eurostat (2026) *Electricity Prices for Household and Non-Household Consumers, H2 2025*. Eurostat.
- [22] Eurostat (2025–2026) *Electricity Price Trends in the EU and Comparison with the United States*. Eurostat.
- [23] European Parliament (2025) *Data Centres and the AI Boom: Implications for Energy Systems*. European Parliament.
- [24] Central Statistics Office Ireland (2025) *Data Centres Metered Electricity Consumption 2024*. CSO Ireland.
- [27, 36] Financial Times (2026) *Mistral's Infrastructure Buildout and European Sovereign AI Ambitions*. Financial Times.
- [29] European Commission (2025) *ReArm Europe / Readiness 2030*. European Commission.
- [30] Bruegel and Kiel Institute for the World Economy (2025) *Defending Europe Without the United States: First Estimates of What Is Needed*. Bruegel / Kiel Institute.
- [31] SIPRI (2026) *Trends in World Military Expenditure 2025*. SIPRI.
- [32] Council of the European Union (2025) *European Defence Spending: 2024 Outturn and 2025 Projection*. Council of the EU.
- [33] NATO (2025) *DIANA: Defence Innovation Accelerator for the North Atlantic*. NATO.
- [35] Reuters (2026) *Thales Inks Deal with Google for Homegrown German Cloud Provider*. Reuters.