RESEARCH STUDY — FEBRUARY 2026

AI FOR AMERICANS FIRST

AI Protectionism, Energy and Semiconductors:
US/Europe Divergence Trajectories 2024–2030

Integrated Geostrategic and Economic Analysis

Chapter II

Fabrice Pizzi
Universite Sorbonne
Master Economic Intelligence — Intelligence Warfare

75% global AI compute = USA $675B US capex 2026 7-12x US/EU ratio

Paris — February 2026

7 chapters • 4 prospective scenarios • 3 geographic zones

Keywords: artificial intelligence, technology protectionism, semiconductors, export controls, sovereign compute, AI geopolitics, France, United States, China

CHAPTER II

Methodology

2.1 General approach: mixed-method multi-scenario prospective analysis

This study combines a retrospective empirical analysis (2020–2026 diagnostic) with a prospective scenario-based projection (2026–2030). This two-pronged architecture responds to the nature of the phenomenon under investigation: AI technology protectionism is simultaneously an observable fact (export controls, Section 232 tariffs) and an ongoing process whose future trajectory depends on discretionary political variables, European strategic responses, and partially unpredictable technological developments.

The retrospective component employs a descriptive quantitative method, based on the aggregation and cross-referencing of data from institutional sources (IEA, SIA/WSTS, Eurostat, EIA), industry reports (McKinsey, Deloitte, Epoch AI), and regulatory documents (Federal Register, BIS, White House). The objective is to establish a rigorous, sourced factual foundation covering three dimensions: data center energy consumption, the semiconductor market, and installed AI compute capacity by region.

The prospective component draws on the scenario planning tradition as formalized by Schwartz (1991) and practiced at Royal Dutch/Shell since the 1970s.[1] This method, belonging to the Intuitive Logics school (Bradfield et al., 2005), consists of constructing plausible and internally consistent scenarios—not to predict the future, but to explore the space of possibilities and evaluate the robustness of different strategies against divergent environmental developments.[2] It is particularly suited to situations characterized by high political and technological uncertainty, where classical econometric models reach their limits—which is precisely the case with AI technology protectionism.

Justification of the methodological choice

Three reasons underpin the choice of scenario methods over pure econometric modelling. First, the key analytical variables are largely political and discretionary: the decision of a US president to impose or not impose GPU quotas on Europe cannot be modelled through a regression function. Second, the interactions between the energy, technological, and geopolitical dimensions are non-linear and systemic: a restriction on GPUs can, through cascade effects, alter energy investment flows, data center location decisions, and the competitive structure of entire sectors. Third, the available data on installed compute by region are partial and heterogeneous: no unified public database of AI FLOPs by country exists, making rigorous econometric calibration premature.

The chosen method therefore combines the quantitative rigour of the empirical diagnostic (sourced data, time series, measurable ratios) with the qualitative flexibility of scenario construction, in the spirit of what Schoemaker (1995) calls a “disciplined heuristic.”[3] The scenarios are not probabilistic forecasts but coherent strategic narratives, each based on explicit assumptions and developing their consequences through measurable metrics.

2.2 Data sources: classification and critical evaluation

The study draws on three categories of sources, whose reliability and potential biases must be explicitly acknowledged. This methodological transparency conforms to the recommendations of the OECD/JRC Handbook on Constructing Composite Indicators (Nardo et al., 2008), which prescribes systematic documentation of sources, their limitations, and their biases in any composite indicator construction.[4]

2.2.1 Primary sources (official and regulatory documents)

This category includes normative or institutional texts: presidential proclamations (Section 232), BIS rules (AI Diffusion Rule, Entity List), IEA reports, European Parliament publications (EPRS), official statistical data (SIA/WSTS for semiconductors, Eurostat and EIA for energy, RTE for France). These sources offer the highest factual reliability but may contain institutional framing biases: the IEA tends to favour moderate scenarios, the European Parliament to emphasize EU sovereignty risks.

2.2.2 Academic sources and think tanks

This category includes peer-reviewed articles (Farrell & Newman, 2019; Bresnahan & Trajtenberg, 1995; Brynjolfsson et al., 2019; Mügge, 2024) and publications from recognized think tanks (Bruegel, Carnegie Endowment, CSIS, OECD, Federal Reserve Board). The former provide robust theoretical grounding; the latter offer empirically founded policy analyses but are potentially influenced by each institution's ideological orientation. We prioritize cross-referencing sources of different orientations (Bruegel / Carnegie / Fed) to limit this bias.

2.2.3 Industry and consulting sources

McKinsey, Deloitte, Accenture, Epoch AI, and CFG Europe provide market data, sectoral projections, and capacity estimates unavailable in public sources. These sources present a potential systematic bias: consulting firms have an interest in amplifying trends (to justify transformation engagements), and market estimates are often optimistic. We mitigate this bias by triangulating figures with institutional data and explicitly flagging discrepancies between sources. For example, 2024 semiconductor sales are $627.6 billion according to the SIA (traditional scope) but $775 billion according to McKinsey (expanded scope)—a 24% gap reflecting methodological differences, not inconsistencies.[5]

2.2.4 AI compute data: the Epoch AI GPU Clusters dataset

Measuring installed AI compute by country constitutes the central methodological challenge of this study. We rely primarily on the Epoch AI GPU Clusters dataset (Pilz, Rahman, Sanders & Heim, 2025), which catalogues over 500 supercomputers and GPU clusters worldwide for the period 2019–2025.[6] This dataset, available under an open Creative Commons Attribution licence, constitutes the most comprehensive and systematically documented source on global

AI compute infrastructure to date. It is used as a reference by the Stanford AI Index Report (2025), by several government reports, and by institutions such as OpenAI and DeepMind.

The dataset covers for each cluster: country of location, chip type (H100, A100, GB200, TPU, etc.), computational performance in 16-bit FLOP/s, number of H100 equivalents, operational date, sector (private/public), electrical power (MW), and estimated hardware cost. This granularity enables aggregation by country and year, directly meeting the needs of our variable F(r) in the CACI.

Limitations of the Epoch AI dataset. Three limitations must be highlighted. First, coverage is estimated at 10–20% of global aggregate AI compute performance (March 2025), with significant heterogeneity across companies and chip types: approximately 20–37% of NVIDIA H100s, 12% of A100s, but less than 4% of Google TPUs and a negligible fraction of custom chips from AWS, Microsoft, or Meta.[7] Second, Chinese systems are anonymized (names removed, values rounded to one significant figure), limiting analytical precision for China. Third, the physical location of a cluster does not determine access: many clusters are accessible via cloud services from other countries.

We complement these data with the OECD Working Paper by Lehdonvirta, Wu, Hawkins et al. (October 2025), which develops a methodology for estimating public cloud AI compute availability by country, counting cloud regions from major providers equipped with AI accelerators (A100, H100, GB200) across 39 economies.[8] This complementary approach distinguishes installed compute (Epoch AI) from accessible compute (OECD), a distinction crucial for the CACI.

2.3 Scenario construction

Scenario construction follows a four-step protocol, inspired by the 2×2 matrix methodology (Schwartz, 1991; van der Heijden, 2004) and adapted to the geostrategic context of AI.[9]

Step 1 — Identification of driving forces

Predetermined elements — whose evolution is reasonably predictable regardless of scenario — include: (i) continued growth in global AI compute demand, (ii) structural increase in data center energy consumption, (iii) European dependence on Asian and American foundries for leading-edge chips, (iv) concentration of the global cloud around three US hyperscalers, and (v) exponential increase in frontier model training costs.

Critical uncertainties — whose evolution depends on political choices, strategic reactions, or technological disruptions — are grouped along two dimensions:

Dimension 1: Intensity of US technology protectionism. This dimension covers a spectrum from maintaining current restrictions to aggressive hardening (GPU quotas for the EU, restrictions on APIs and models, explicit prioritization of deliveries to US companies).

Dimension 2: European response capacity. This dimension covers a spectrum from passive posture (marginal adaptation, acceptance of dependence) to active response (Compute Zones with derogated energy, accelerated AI Factories, SMR nuclear for data centers, alternative partnerships Japan-Korea-Taiwan, AI Act revision).

Step 2 — 2×2 matrix and scenario generation

Crossing the two uncertainty dimensions generates a four-scenario matrix. The choice of four rather than three scenarios is deliberate. Schwartz (1991) and Shell method practitioners recommend never building three scenarios, as the human mind tends to treat the middle scenario as the "most likely," thereby reducing the exercise's utility.[10] The 2×2 matrix forces the analyst to explore extreme quadrants—precisely where strategic ruptures play out.

Step 3 — Narrative development and quantification

Each scenario is developed following a standardized protocol comprising three components: (a) a strategic narrative describing the plausible sequence of events between 2026 and 2030, (b) a quantification of key metrics calibrated on 2020–2026 empirical data and projected according to the scenario's assumptions, and (c) early warning indicators (leading indicators) enabling identification, from 2026–2027 onward, of which scenario reality is converging toward.

Step 4 — Sensitivity analysis and robustness

For each recommendation formulated in Chapter VII, we evaluate its robustness across all four scenarios. A recommendation is considered robust if it produces positive or neutral results in at least three of the four scenarios.

2.4 Key metrics and original indicator: the CACI

2.4.1 The six divergence metrics

We define six metrics to be calculated or estimated in the empirical diagnostic (Chapter III), then projected in each scenario (Chapter V). Together, they form a dashboard of US/EU divergence in AI: M1 (compute gap: US/EU installed AI FLOPs ratio normalized by GDP), M2 (relative FLOP cost for training), M3 (cloud dependence: share of EU AI workloads on US infrastructure), M4 (sectoral AI productivity), M5 (energy constraint: data center demand/capacity ratio), and M6 (AI relocations).

2.4.2 The Compute-Adjusted Competitiveness Index (CACI): theoretical foundations and construction

Grounding in the literature. The construction of a compute-centred AI competitiveness composite indicator responds to a need identified by several converging research streams. Since 2023–2024, academic and institutional literature increasingly emphasizes that computational capacity has become the most discriminating factor of production for frontier AI (Sevilla et al., 2022; Epoch AI, 2025; Pilz et al., 2025). US export controls (BIS, October 2022; updated 2023 and 2025) explicitly place advanced compute at the heart of geopolitical competition, while Hawkins, Lehdonvirta & Wu (2025) introduce the concept of "compute sovereignty" as a structuring dimension of strategic autonomy.[11]

Yet existing AI competitiveness indices do not place compute at the centre of their construction. The IMF's AI Preparedness Index (Cazzaniga et al., 2024), covering 174 countries, aggregates four dimensions (digital infrastructure, human capital, innovation/economic integration, regulation/ethics) without directly measuring installed compute capacity.[12] The Tortoise Media Global AI Index (2024), ranking 83 countries on 122 indicators across three pillars (Implementation, Innovation, Investment), includes an infrastructure/supercomputing component but subsumes it within a weighted additive index where compute is merely one factor among many.[13] The Stanford AI Index (2025) provides rich compute trend data but proposes no composite index. None of these instruments formalizes the mechanism by which compute, adjusted for its energy cost and related to an economy's absorptive capacity, determines AI competitiveness.

The CACI aims to fill this gap by proposing a parsimonious yet theoretically grounded indicator that captures the multiplicative interaction between three factors: accessible compute, the energy cost constraining it, and the economic and human capacity to exploit it.

Formal definition. The CACI for a region r at period t is defined as:

$$\mathbf{CACI(r,t) = [\ F(r,t) \times E(r,t)^{-1}\ ] \ / \ [\ GDP(r,t) \times L(r,t)\ ]}$$

where:

F(r,t) = installed and accessible AI compute capacity in region r at period t, measured in aggregate FLOP/s (16-bit performance). We use Epoch AI GPU Clusters data aggregated by country and year, complemented by OECD estimates of accessible cloud compute. The unit is PetaFLOP/s (10^{15} FLOP/s), including domestic cloud capacities and authorized access quotas to foreign clouds.

E(r,t) = average energy cost of compute in region r, in €/MWh for data centers. This factor adjusts raw compute for its energy constraint: at equal FLOPs, a country with electricity twice as expensive has a CACI twice as low. Data come from Eurostat (industrial tariffs by country) and the EIA (US prices), adjusted for hyperscaler Power Purchase Agreements via IEA (2025) estimates.

GDP(r,t) = gross domestic product of region r (World Bank, Eurostat). GDP normalization ensures comparability across economies of very different sizes: without it, the United States and China would mechanically dominate any ranking by sheer economic mass.

L(r,t) = working population with AI competencies (proxy: STEM graduates + certified AI training). This factor captures absorptive capacity in the sense of Cohen & Levinthal (1990): abundant compute without human capital to exploit it does not produce competitiveness. OECD data on STEM graduates are complemented by LinkedIn Economic Graph estimates on AI skills density by country.[14]

Justification of the multiplicative form. The choice of multiplicative (geometric) rather than additive (arithmetic) aggregation is deliberate and rests on three arguments. First, General Purpose Technology theory (Bresnahan & Trajtenberg, 1995) posits strong complementarity between innovation inputs: compute without affordable energy or qualified human capital produces no competitiveness gains, justifying a form where weakness in one factor penalizes the whole. Second, the OECD/JRC Handbook (2008, p. 33) recommends geometric aggregation when components are not perfectly substitutable: unlike the arithmetic mean, the geometric mean does not allow a very high score on one dimension to fully compensate a very low score on another. Third, recent work on AI index construction (Koronakos, Kritikos & Sotiros, 2024; analysis of the Tortoise GAII via Choquet integral) confirms that AI competitiveness dimensions exhibit interactions (complementarities and redundancies) that make simple linear aggregation problematic.[15]

Economic interpretation. In logarithmic form, the CACI can be written: ln(CACI) = ln(F) − ln(E) − ln(GDP) − ln(L). This transformation is convenient for econometric analysis as it linearizes the relationship and allows each coefficient to be interpreted as an elasticity. The indicator is designed for bilateral comparison: the ratio CACI(US)/CACI(EU) or CACI(US)/CACI(FR) measures relative competitive advantage. A ratio of 7 means that, per unit of GDP and at equal human capital, US actors have seven times more effective compute (adjusted for energy cost) than European actors.

2.4.3 Calibration protocol and data sources

CACI calibration follows a four-step protocol, consistent with OECD/JRC Handbook (2008) recommendations for composite indicator construction: (i) identification and collection of raw data, (ii) treatment of missing values and normalization, (iii) aggregation, and (iv) sensitivity analysis.

F(r,t) — Installed AI compute capacity. Estimation follows a bottom-up approach in three layers. The main layer comes from the Epoch AI GPU Clusters dataset (February 2026 version, 746 clusters), providing 16-bit FLOP/s performance aggregated by country for 2019–2025. National shares as of May 2025 are: United States ~74.5%, China ~14.1%, European Union ~4.8%, Norway ~1.8%, Japan ~1.4%.[16] The complementary layer comes from the OECD (Lehdonvirta et al., 2025), cataloguing cloud regions with AI accelerators across 39 economies, capturing the "accessibility" dimension that physically installed compute does not fully reflect. The third layer uses partial data published by hyperscalers (Microsoft, Google, Meta, OVHcloud) and CFG Europe estimates for European capacity, cross-referenced with the Top500 ranking for public HPC.

F aggregation procedure. For each country-year, we: (1) extract from the Epoch AI dataset the sum of 16-bit FLOP/s from all operational clusters located in the country, filtering for confirmed systems (certainty ≥ "Likely"); (2) apply an extrapolation factor to correct for the dataset's under-coverage (~10–20% of global compute), using Epoch AI's sectoral coverage estimates by chip type; (3) add estimated accessible cloud capacities via the OECD methodology for countries where foreign cloud compute represents a significant share of effective capacity. Raw data and calculation code (Python/pandas) are documented in the methodological appendix.[17]

E(r,t) — Energy cost. Eurostat data (industrial electricity tariffs by country, consumption band IE) and EIA data (Average Retail Price of Electricity, Industrial) provide the foundation. We adjust for negotiated large-consumer tariffs (hyperscaler PPAs), using IEA (2025, Energy and AI) estimates of data center energy mix by region. EU costs are typically 2 to 3 times US levels for industrial electricity, before PPAs.[18] The Federal Reserve Board (October 2025) documents a significant negative correlation between energy costs and AI adoption at the European firm level.

GDP(r,t) and L(r,t). GDP is available from the World Bank (World Development Indicators) and Eurostat. The AI human capital proxy L(r,t) combines three sub-indicators: (i) number of STEM graduates (OECD, Education at a Glance), (ii) AI skills density as measured by LinkedIn Economic Graph (profiles with AI skills relative to working population), and (iii) AI certifications (estimates based on certified AWS/Google/Microsoft cloud programs by country). This proxy is consistent with approaches used by the Federal Reserve Board (2025) and the IMF in the AI Preparedness Index (Human Capital component). We acknowledge a bias favouring English-speaking countries and economies where LinkedIn is dominant, and document this bias's effect in the sensitivity analysis (section 2.4.5).[19]

2.4.4 Positioning relative to existing AI competitiveness indices

The IMF AI Preparedness Index (AIPI) covers 174 countries (2023) and aggregates four pillars: digital infrastructure, human capital and labour market policies, innovation and economic integration, regulation and ethics. The AIPI does not measure installed AI compute and does not weight by energy cost. Its expected correlation with the CACI is positive but imperfect: countries scoring highly on the AIPI (Singapore, Denmark, Netherlands) are not necessarily those with the most effective compute per unit of GDP.[20]

The Tortoise Media Global AI Index (GAII) ranks 83 countries on 122 indicators across three pillars (Implementation, Innovation, Investment). It includes an infrastructure/supercomputing component but aggregates it linearly with subjective weights (acknowledged by Tortoise as a limitation). The GAII is broader and more multidimensional than the CACI, but precisely because it is broad, it dilutes the compute signal in a many-indicator composite. As Koronakos et al. (2024) show, the GAII's weighting subjectivity can invert country rankings depending on chosen weighting scenarios.

The Stanford AI Index (2025) constitutes the most comprehensive reference in terms of raw data (notable models by country, investment, publications, patents, compute trends). It does not propose a composite index but provides the time series used by many other indices. The Stanford AI Index's public data are available via an open Google Drive folder.[21]

CACI's specific value added. The CACI distinguishes itself through four properties: (i) it places compute at the centre rather than the periphery of the indicator, reflecting compute's now-dominant role in frontier AI; (ii) it explicitly integrates energy cost as a bottleneck, consistent with IEA (2025) findings; (iii) it uses a theoretically grounded multiplicative aggregation rather than additive; and (iv) it is parsimonious (four variables), making it transparent and reproducible, at the cost of lesser comprehensiveness.

2.4.5 CACI limitations and sensitivity analysis

First limitation: opacity of F(r,t). Installed compute measurement depends on incomplete private data. The Epoch AI dataset covers only 10–20% of global compute, with uneven coverage across sectors and companies. National attribution is itself debatable: a growing share of compute is held by a few private hyperscalers operating globally. Mitigation: we systematically document uncertainty margins, present ranges rather than point values, and verify result stability when F varies by ±30%.

Second limitation: qualitative heterogeneity of compute. The CACI aggregates FLOPs without distinguishing GPU generations (an H200 does not equal an A100 in energy efficiency and real performance). Mitigation: the Epoch AI dataset provides performance in H100 equivalents, offering partial normalization. We propose a GPU generation weighting factor in the appendix and show that adjustment marginally modifies rankings.

Third limitation: the human capital proxy L(r,t). The STEM + LinkedIn + certifications combination presents a bias favouring English-speaking countries. This bias likely underestimates China and some Asian economies. Mitigation: we replicate the analysis using the IMF AIPI's Human Capital sub-index as an alternative proxy and show ranking sensitivity to this choice.

Fourth limitation: indicator staticity. The CACI measures a state at a given moment, not a dynamic. This is why we calculate it across multiple years (2022, 2024, 2026) and project it in each scenario, enabling trajectory tracking and gap evolution assessment.

Fifth limitation: endogeneity. Countries gaining AI productivity invest massively in compute, creating a reverse causality risk: the CACI might capture the consequence rather than the cause of competitiveness. This study, within its dissertation framework, does not formally instrument this relationship (no instrumental variables or GMM strategy). However, we note that the exogenous shock of the October 2022 BIS rules offers a natural quasi-experiment that could ground a causal identification strategy in Difference-in-Differences: China (treated) experiences an abrupt ceiling on F while the United States (control) accelerates. Our Chapter III figures illustrate this divergence. We identify the formal treatment of this endogeneity (DiD, IV, or Arellano-Bond GMM) as a priority research avenue for any publishable extension of this work.[22]

Sixth limitation: normalization bias for small economies ("small-economy bias")

Comparative analysis of the CACI against the IMF AIPI and Tortoise GAII indices reveals an instructive anomaly: certain countries with low GDP and a reduced AI workforce obtain disproportionately high CACI scores, radically out of step with their rankings on other indices. South Africa is emblematic: ranked last or second-to-last on the IMF AIPI and Tortoise indices, it appears as the global leader on the CACI (normalized score of 100, surpassing the United States).

Mechanical explanation. This counter-intuitive result is entirely explained by the multiplicative structure of the denominator GDP(r) × L(r). South Africa has a relatively modest GDP (~$400 billion) and a very small AI workforce (L). The CACI denominator is therefore minuscule. Meanwhile, the country has some compute clusters (non-zero numerator F) and industrial electricity among the cheapest in the world (Eskom tariffs ~$0.05–0.07/kWh, despite load-shedding episodes), keeping E^{-1} at a high level. Dividing a modest but non-negligible numerator by a very small denominator mechanically produces an artificially inflated score.

Methodological diagnosis. This phenomenon is a classic case of *normalization bias* identified in the composite indicator literature. The OECD/JRC Handbook (Nardo et al., 2008, pp. 27-

29) warns that normalization by GDP or population can produce misleading results for small economies: just as Luxembourg's or Iceland's GDP per capita does not reflect "productivity" superior to the United States, South Africa's CACI does not reflect superior AI competitiveness. The problem is structural: in the formula CACI = $[F \times E^{-1}] / [GDP \times L]$, a country can achieve a high score without critical mass of compute, simply because its denominator is small enough.[25]

Correction strategies. We propose three mitigation mechanisms, in order of increasing intervention:

(a) Critical mass threshold on F. Exclude from CACI calculation countries whose installed compute capacity F(r) falls below a minimum threshold (e.g., the top 20 countries in absolute FLOPs from the Epoch AI dataset, or an absolute threshold of 10,000 H100 equivalents). This threshold ensures the CACI is only calculated for economies that have reached sufficient critical mass for the indicator to be meaningful.

(b) Scaling factor. Introduce a multiplicative factor linked to absolute compute mass: $CACI^*(r) = CACI(r) \times \alpha(r)$, where $\alpha(r) = \min(1, F(r)/F\_median)$. This factor progressively penalizes countries whose absolute compute falls below the global median, avoiding the discontinuities of a binary threshold while correcting the bias.

(c) Dual ranking (intensity vs. scale). Systematically present the CACI in two forms: an *intensity* ranking (CACI as defined, suited for bilateral US/EU/FR comparisons) and a *scale* ranking (raw F, suited for absolute power comparisons). The Tortoise index (2024) already uses this distinction between AI capacity (scale) and AI capacity (intensity). The CACI as defined corresponds to an intensity measure and should be interpreted as such.

In the present study, we primarily adopt solution (a): the CACI is calculated and discussed for countries exceeding the critical mass threshold, and bilateral comparisons focus on major economies (United States, EU, France, China). Results for smaller economies (Brazil, South Africa, UAE, etc.) are presented as indicative, with the explicit caveat that normalization bias affects them. This transparency conforms to the OECD/JRC Handbook recommendation to always document cases where the indicator produces potentially misleading results.[26]

Despite these limitations, the CACI addresses a need identified in the literature: no formalized compute-adjusted competitiveness indicator currently exists for systematic regional comparison. The methodological contribution lies in the framework rather than in the precision of figures: even with approximate data, the CACI makes visible the structural gap that traditional indicators (GDP, R&D spending, patents) do not capture.

2.5 Scope and delimitations

Geographic scope. The analysis focuses on the bilateral United States / European Union relationship, with a specific focus on France. China is treated as a contextual variable (primary target of US export controls, factor of pressure on chip production capacities) but is not subject to in-depth analysis. Japan, South Korea, and Taiwan appear as semiconductor supply chain actors.

Temporal scope. The diagnostic covers 2020–2026, scenarios cover 2026–2030. The 2030 horizon is chosen as it corresponds to the convergence of several deadlines: IEA projections for data center energy, expected maturity of the EU Chips Act, France 2030 SNIA objectives, and potential arrival of the first operational SMR nuclear reactors.

Technological scope. The study covers frontier AI (foundation models, compute-intensive) and its material prerequisites (GPU/ASIC, data centers, energy). It integrates AI robotics as an amplifying factor for energy demand. It does not address edge embedded AI (smartphones, IoT), except insofar as it constitutes a specific objective of the French SNIA.

2.6 General methodological limitations

Radical political uncertainty. Technology protectionism depends on discretionary political decisions with structurally low predictability. A change of US administration in 2028, an unexpected US-EU trade agreement, or an escalation of the US-China conflict could invalidate certain assumptions. This is precisely why we propose four scenarios rather than a single trajectory.

Technological disruptions. The DeepSeek episode (January 2025), where a Chinese model achieved near-frontier performance with substantially reduced training budget, illustrates the possibility of efficiency breakthroughs that would alter the problem's terms. The IEA (2025, Energy and AI) devotes a case study to DeepSeek and concludes that even with significant efficiency improvements, demand growth absorbs gains (Jevons rebound effect).[23]

Compute data opacity. The exact number of GPUs deployed per hyperscaler, the precise geographic distribution of data centers, and GPU volumes exported by region are partially or fully confidential data. Our installed compute estimates carry significant margins of error, which we systematically document.

Consulting source bias. As noted in section 2.2, industry sources carry a systematic optimism bias. We mitigate this through triangulation but cannot eliminate it entirely.

These limitations do not compromise the analysis's validity. The scenario method is precisely designed to function in high-uncertainty environments, where the objective is not prediction but structured exploration of possibilities. As Schwartz notes, "scenarios are not forecasts; they are plausible stories that help you think."[24] Our contribution lies in the rigour of the framing, the explicitness of assumptions, transparency of data sources, and the originality of the CACI indicator, rather than in the precision of numerical projections.

Notes

[1] Schwartz, P. (1991), The Art of the Long View: Planning for the Future in an Uncertain World, New York, Doubleday. See also Wack, P. (1985), "Scenarios: Uncharted Waters Ahead," Harvard Business Review, Sept.–Oct. 1985, pp. 72-89.

[2] Bradfield, R., Wright, G., Burt, G., Cairns, G. & Van Der Heijden, K. (2005), “The Origins and Evolution of Scenario Techniques in Long Range Business Planning,” Futures, 37(8), pp. 795-812.

3 Schoemaker, P.J.H. (1995), “Scenario Planning: A Tool for Strategic Thinking,” MIT Sloan Management Review, 36(2), pp. 25-40.

4 Nardo, M., Saisana, M., Saltelli, A. & Tarantola, S. (2008), Handbook on Constructing Composite Indicators: Methodology and User Guide, OECD Publishing, Paris.

5 The SIA/McKinsey gap is explained by scope: McKinsey (January 2026, “Hiding in Plain Sight”) includes the value of captive designers (Apple, Amazon, Tesla) and fabless operators whose sales do not appear in WSTS statistics.

6 Pilz, K.F., Rahman, R., Sanders, J. & Heim, L. (2025), “Trends in AI Supercomputers,” arXiv:2504.16026, April 2025. Dataset accessible at https://epoch.ai/data/gpu-clusters under Creative Commons Attribution licence.

7 Epoch AI (2025), GPU Clusters Documentation. Coverage estimated at ~20–37% of NVIDIA H100s, ~12% of A100s, and ~18% of AMD MI300X, but less than 4% of Google TPUs.

8 Lehdonvirta, V., Wu, B., Hawkins, Z.J., Caira, C. & Russo, L. (2025), “Measuring Domestic Public Cloud Compute Availability for Artificial Intelligence,” OECD AI Papers, No. 49, https://doi.org/10.1787/8602a322-en.

9 Van der Heijden, K. (2004), Scenarios: The Art of Strategic Conversation, 2nd ed., Chichester, Wiley.

10 Schwartz (1991), op. cit., pp. 241-243.

11 Hawkins, Z.J., Lehdonvirta, V. & Wu, B. (2025), “AI Compute Sovereignty,” SSRN, June 2025, https://ssrn.com/abstract=5312977. Sevilla, J. et al. (2022), “Compute Trends Across Three Eras of Machine Learning,” arXiv:2202.05924.

12 Cazzaniga, M. et al. (2024), “Gen-AI: Artificial Intelligence and the Future of Work,” IMF Staff Discussion Note 2024/001. The AI Preparedness Index is accessible via the IMF dashboard: https://www.imf.org/external/datamapper/AI_PI@AIPI/.

13 Tortoise Media (2024), The Global Artificial Intelligence Index 2024. Methodology available at https://www.tortoisemedia.com/data/global-ai. See also Koronakos, G., Kritikos, M. & Sotiros, D. (2024), “Mitigating Subjectivity and Bias in AI Development Indices,” Expert Systems with Applications.

14 Cohen, W.M. & Levinthal, D.A. (1990), “Absorptive Capacity: A New Perspective on Learning and Innovation,” Administrative Science Quarterly, 35(1), pp. 128-152.

15 Geometric aggregation is also used by the UNDP Human Development Index since 2010, for analogous reasons of non-substitutability between dimensions. See OECD/JRC Handbook (2008), pp. 31-33.

16 Epoch AI (2025), “The US Hosts the Majority of GPU Cluster Performance, Followed by China,” June 2025. Data at https://epoch.ai/data-insights/ai-supercomputers-performance-share-by-country.

17 The Python code and data are reproducible. The Epoch AI dataset is downloadable in CSV at https://epoch.ai/data/gpu_clusters.csv (daily refresh). Full documentation at https://epoch.ai/data/gpu-clusters-documentation.

18 Federal Reserve Board (October 2025) documents a significant negative correlation between energy costs and AI adoption at the European firm level. IEA (2025), Energy and AI.

19 The LinkedIn bias is documented: in countries where LinkedIn is rarely used (China, Russia, some Southeast Asian economies), AI skills density is mechanically underestimated.

20 IMF (2024), AI Preparedness Index Dashboard, https://www.imf.org/external/datamapper/datasets/AIPI. Singapore, Denmark, Netherlands, and the US occupy top positions; China ranks 31st (score 0.63).

21 Maslej, N. et al. (2025), “Artificial Intelligence Index Report 2025,” Stanford Institute for Human-Centered AI, April 2025. Public data: https://drive.google.com/drive/folders/1AxxxL9-AsaeMdDKtTNHCR1KqEJTsHCod.

22 The most promising causal identification strategy would exploit the exogenous shock of the October 2022 BIS rules in a Difference-in-Differences (DiD) framework. See also RAND (2025). For GMM, see Arellano, M. & Bond, S. (1991), “Some Tests of Specification for Panel Data,” Review of Economic Studies, 58(2), pp. 277-297.

23 IEA (2025), Energy and AI, devotes a case study to DeepSeek and concludes that even with significant efficiency improvements, demand growth absorbs gains (Jevons rebound effect).

24 Schwartz (1991), op. cit., p. 38. Author’s translation.

25 The problem is analogous to the “Singapore effect” in competitiveness indices: small, open, specialized economies (Singapore, Luxembourg, Ireland) systematically dominate GDP- or population-normalized rankings,

not because they are structurally superior, but because the denominator amplifies the signal. See Saisana, M. & Tarantola, S. (2002), State-of-the-art Report on Current Methodologies and Practices for Composite Indicator Development, JRC European Commission, pp. 14-16.

26 OECD/JRC Handbook (2008), pp. 27-29: “When normalising by GDP or population, users should be aware that small economies may obtain extreme values [...] Presenting both raw and normalised data is recommended.” The dual ranking solution (intensity/scale) is also used by Tortoise Media (2024) in the Global AI Index.