

The Electricity Supply Bottleneck on U.S. AI Dominance

By Cy McGeady, Joseph Majkut, Barath Harithas, and Karl Smith

It is now well understood that the rapid technological progress of artificial intelligence (AI) has profound energy sector implications. AI technology is effectively the result of three inputs: chips, data, and electricity. This paper focuses on electricity on the basic premise that electricity supply is the most acutely binding constraint on expanded U.S. computational capacity and, therefore, U.S. AI dominance.

This paper starts with a survey of demand-side forecasts. It then highlights data on the geographic distribution of data center development currently underway in the United States, the supply-side dynamics underway in response to demand growth, and challenges to meeting this new demand. The role of coal, gas, renewables, and nuclear power in meeting new demand are each assessed. The central principle for understanding these developments is speed-to-power, or the measure of how fast a potential data center site can access the electricity needed to power its stock of chips.

Speed-to-power should likewise be used to organize federal policymakers' approach to permitting policy and use of emergency authorities in the near term. On the other hand, five years from now is tomorrow in the power sector. A severe near-term supply crunch must not distract policymakers from the need for long-term thinking in the electricity sector. Numerous long-standing policy challenges in the power sector deserve renewed attention, including gas-electric coordination, interregional seams management, and improved cost efficiency in transmission planning. This paper closes by proposing several new policies and authorities that contribute to these issues, but which are primarily organized around establishing U.S. electricity supply dominance in a bid to advance U.S. AI leadership. A new era of electricity-intensive economic growth has arrived, and the need for strategic thinking in the electricity sector has never been greater.

The Age of AI and Electricity Demand

The basic reality of a surge in data center-based electricity demand has been confirmed by a wide range of work from the private sector, civil society, and national labs. Recent **estimates** from the Lawrence Berkeley National Laboratory (LBNL) place electricity consumption by data centers at 176 terawatt-hours (TWh) in 2023, representing 4.4 percent of total U.S. electricity demand.

Table 1: Data Center Electricity Consumption Forecasts

Group	Forecast Subject	Current Value	Forecast Value	Growth Value
CSIS (author's analysis)	U.S. AI data centers	4 gigawatts (GW) in 2024	84 GW by 2030	2,100 percent
LBNL	U.S. data centers	20 GW in 2023	74 to 132 GW by 2028	370–660 percent
RAND	Global AI data centers	11 GW in 2024	68 GW by 2027 and 327 GW by 2030	618 percent
SemiAnalysis	Global data centers	49 GW in 2023	96 GW by 2026	196 percent
BCG	Global data centers	60 GW in 2023	127 GW by 2028	212 percent
McKinsey	Global data centers	55 GW in 2023	171 to 219 GW by 2030	311–398 percent
Goldman Sachs	Global data centers (excluding crypto)	400 TWh in 2023	1,040 TWh by 2030	260 percent

Source: Authors' analysis.

The range in future estimates of AI power demand highlights the complex set of factors—hardware technology, algorithmic progress, commercial strategy, economy-wide uptake of AI, and power sector capacity—that interact to create uncertainty over the exact trajectory of electricity demand from the computation sector. The sector is attracting enormous volumes of **capital investment** and competition is driving rapid innovation throughout the ecosystem. Developments including sudden efficiency jumps such as those achieved by **DeepSeek** or progress on **distributed data center training** capabilities are to be expected. Such developments will impact specific firms, commercial strategies, and technology paths, but are indicative of continued sectoral scaling rather than signs of imminent sectoral crash correction. Policy should see past short-term perturbations and grasp that growth is the definitive long-term direction of AI technology and computation demand.

Despite the dynamic nature of the current moment, policymakers can be certain a new era of electricity demand growth has arrived. Data from SemiAnalysis, which provides best-in-industry tracking of chip production, chip orders, and individual data center developments, shows that **over 80 GW of data center capacity** under various stages of development could be brought online in the United States by 2030. These facilities could consume over 800 TWh per year, which alone represents a 3 percent annual growth in total U.S. power demand. The biggest risk to this forecast is in the electric power sector's ability to serve this demand.

The U.S. electric power sector is facing a stunning and sudden paradigm shift. For roughly two decades, top-line national electricity consumption has **stagnated**, growing at a compound annual growth rate of nearly 0 percent since 2007. The electric power industry as a whole has been decelerating since the 1970s; recent decades of near-zero demand growth follow decades more of steadily declining growth rates. Multiple generations of commercial strategy, regulatory norms, and policy debates have been conditioned by this seemingly inexorable trajectory and are now out of date.

This **story** extends beyond AI. Electricity demand is also growing from other electricity-intensive industries like semiconductor fabrication and battery manufacturing. The broad political consensus to reindustrialize the U.S. economy will drive growth in energy-intensive industries like mining, minerals processing, metallurgy, and beyond. A deep technological trend toward electrification means industry, along with the transport and heating sectors, is growing more electricity-intensive each year. Successfully navigating a new era of electricity demand growth will deliver the United States a lasting advantageous position in the **technological commanding heights** of the future.

The Future of Data Center Demand

Today, access to electricity supply is the binding constraint on expanded computational capacity and therefore on continued U.S. leadership in AI. This fact is demonstrated by a total focus among data center developers on “speed-to-power.” Speed-to-power is the time it takes a potential data center site to receive access to electricity supply. In Northern Virginia—the nation's and the world's largest data center market—speed-to-power is growing worse, as data centers now face electricity supply **wait times** up to 7 years.

For data center developers, speed-to-power far outweighs other factors like the price of power or access to land. Even access to high-end chips is a secondary concern, as hyperscalers cannot access enough electricity supply to power their existing stocks of chips. An example of the high value placed on speed-to-power relative to price is the xAI **data center facility** in Memphis, Tennessee, which, due to long wait times for grid-supplied power, instead rented road-portable gas-fired generators which operate at far higher unit costs than large grid-connected combined cycle power plants. The race for progress on the AI frontier and the rapid growth in computational demand for AI services make speed-to-power the central principle driving data center investment in the near term.

Data from SemiAnalysis again provides **clear indications on the scale and distribution** of this demand boom. Virginia, already the world's largest data center market, is on track to see enormous growth over the next five years and will remain the country's most important computing cluster. Despite severely constrained power supply data centers continue to expand in the area **because** of access to key internet infrastructure such as fiber networks, latency, and other provision of service

considerations. By 2030, the region could **host** 20 GW of data center capacity. A central policy objective for federal AI strategy should be to improve speed-to-power for this computing cluster.

Table 2: Data Center Boom by Region (Active and Planned Capacity, GW)

Region	2024	2030	Growth
Southeast	13	49	36
West	12	43	31
Midwest	7	30	23
Northeast	1	3	2
U.S. Total	32	124	92

Source: “Datacenter Industry Model,” SemiAnalysis, February 2025, <https://semianalysis.com/datacenter-industry-model/>.

Those data centers that are not constrained by service provision concerns, such as data centers dedicated to model training, are seeking out new geographies that offer faster speed-to-power. SemiAnalysis shows that **twenty-nine states** are slated to see over 100 percent growth in hosted data center capacity. States like **Louisiana** and **Mississippi** currently host no data center capacity and have recently attracted multi-gigawatt datacenter investments. Despite a flight to new geographies, computing capacity will remain regionally concentrated. By 2030 just nine states will host 70 percent of the nation’s data center capacity. Virginia and Texas are the standouts, projected to together represent 34 percent of the nation’s data center capacity in 2030.

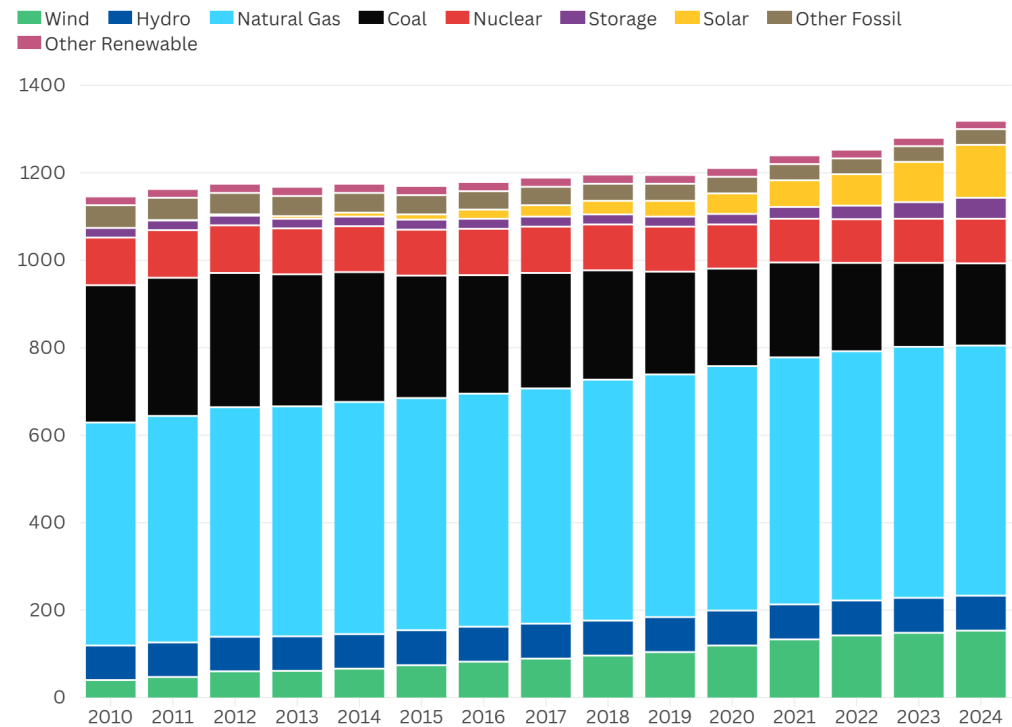
Texas, the Midwest, Southeast, and Southwest stand out as new regions attracting large volumes of data center investment. In contrast, California and the Northeast stand out for low levels of data center development. Data center investment is flowing where state-level power sector policy, permitting, and land-use issues are permissive to a rapid buildout of new generation needed to power new data centers.

The State of National Generation Base

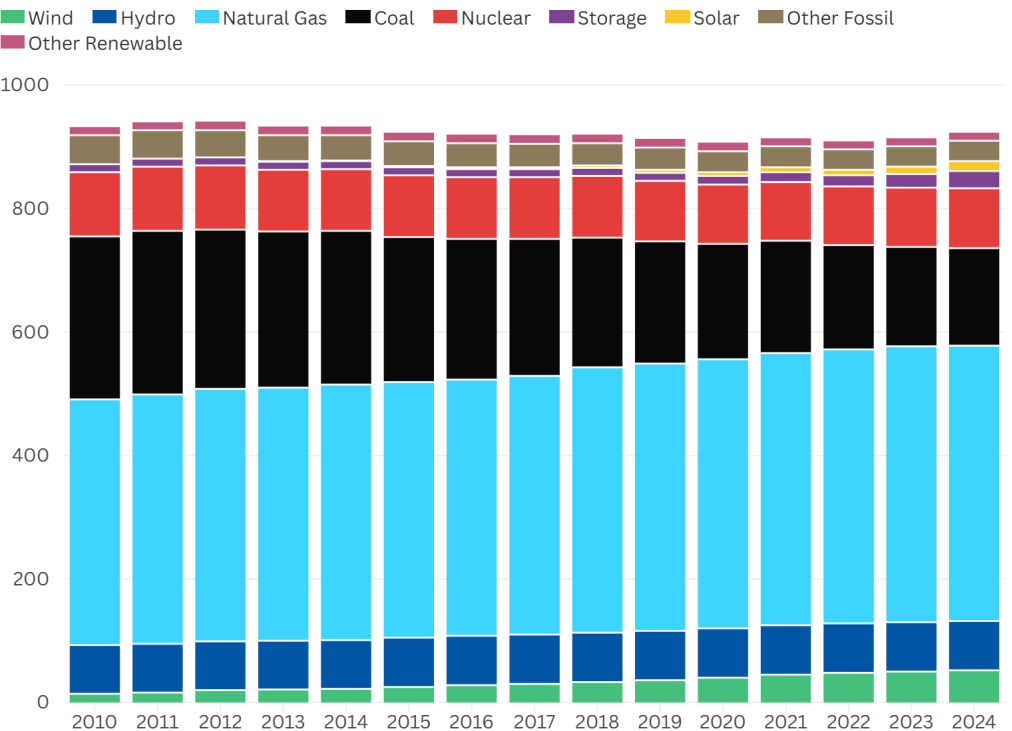
How did electricity supply suddenly emerge as a binding constraint on data center expansion and AI progress in 2025? After all, Energy Information Administration (EIA) **data** shows that since 2010, nameplate generation capacity in the United States has grown by 172 GW to a total of 1,318 GW.

The non-firm nature of wind and solar generation makes nameplate capacity a deceptive measure of the nation’s generation base. To maintain reliability, utilities and grid operators account for and plan using the effective capacity of generation resources, which accounts for the likely availability of each class of generation technology during peak demand scenarios. An example is the **Effective Load Carrying Capacity** (ELCC) measures used by PJM, the largest power grid in the nation, in its capacity markets, which are designed to ensure sufficient generation resources to meet demand over the long term. Applying the PJM ELCC factors to the nameplate capacity dataset results in a dramatically different picture of the national generation mix.

Figure 1: U.S. Generation Mix-Nameplate Capacity vs. Effective Capacity (GW)



Source: Author's calculations; Preliminary Monthly Electric Generator Inventory, U.S. Energy Information Administration, February 2025, <https://www.eia.gov/electricity/data/eia860m/>.



Source: Author's calculations; Preliminary Monthly Electric Generator Inventory, U.S. Energy Information Administration, February 2025, <https://www.eia.gov/electricity/data/eia860m/>.

Though this is a rough adjustment—in reality each utility and grid operator employs distinct capacity adjustment factors—the overall effect is directionally correct. The total effective capacity of the U.S. generation base has stagnated since 2010, and it may have even declined. Coal-fired generation with high ELCC ratings (84 percent) has been replaced by low ELCC resources like onshore wind (34 percent) and solar (13 percent). Even dispatchable gas-fired generation (78 percent) has a lower rating than coal and nuclear (95 percent) because of fuel supply and gas-electric coordination **issues** during winter storms.

A stagnant base of effective capacity has only been possible (i.e., compatible with the reliability imperative) because it coincided with a period of near-zero load growth at the national level. And yet, even small amounts of demand growth combined with a flat or declining base of effective capacity equates to thinning reserve margins. This is a finding compatible with **repeated reporting** from the North American Electric Reliability Corporation (NERC) and reports from regional grid operators like **MISO** and **PJM**, which all warn of thinning generation reserve margins. A series of capacity shortfall incidents in California (**2020, heat wave**, Western Interconnection), Texas (**2021, Winter Storm Uri**, Electric Reliability Council of Texas (ERCOT) Interconnection), and the southeast (**2022, Winter Storm Elliot**, Eastern Interconnection) have demonstrated that increasingly narrow capacity margins are leading to reliability failures.

On a national level, there is effectively no “spare capacity.” Though regional pockets and individual generators where spare capacity exists, these are exceptions to the broader national trend. Today, every new gigawatt of data center demand must be met with matching new gigawatts of effective capacity sited within the borders of the same reliability planning region. The past failure to grow effective capacity explains why a focus on speed-to-power necessarily follows from the data center boom and AI technology race.

The Coal Option

The sudden emergence of electricity demand growth has definitively slowed the rate of decline in the coal fleet. Major utilities have **proposed** integrated resource plans (IRPs) with **suspended or delayed** coal retirement schedules. **Soaring capacity prices** in PJM have improved prospects for merchant-owned coal plants. Rising market **valuations** for coal plant operators further illustrate the improved economic outlook for existing coal.

As of December 2024, the U.S. coal fleet is **composed** of over 400 units representing 188 GW of capacity. As recently as 2023, **expectations** were for 70–100 GW of this capacity to retire by 2035. The Environmental Protection Agency’s (EPA) modeling for its 2024 **greenhouse gas emissions rule** indicated 150 GW or more of retirements were possible by 2035. But the Trump administration’s goal to **repeal** the EPA greenhouse gas rule, combined with improving market signals and shifting utility IRPs, means these rapid retirement scenarios are unlikely to materialize. In fact, the Trump administration’s exploration of using emergency authorities to keep coal plants open is unlikely to be broadly necessary.

In the near-term speed-to-power era, delayed coal retirements make the problem of supplying new AI demand more manageable. The retirement of a coal plant creates a “backfill” requirement for new generation that delivers the same amount of effective generation capacity. Preserving reliability is the first priority for utilities and reliability authorities, so new generation capacity is generally allocated to

the backfill requirement before new demand customers like data centers. In short, backfill competes with new demand for a limited supply of new generation projects and must always win. Therefore, slowed retirement schedules mean that most new generation resources can be allocated to serve new data center demand, a result which increases speed-to-power for AI data centers.

Improved near-term prospects notwithstanding, the coal fleet is aging and remains in terminal decline. Over 130 GW of the capacity (70 percent of the fleet) is at least 40 years old. Age and declining economic competitiveness with gas and renewables has pushed down utilization; in 2023 the coal fleet nationwide produced at a 42 percent **capacity factor**, down from 61 percent in 2014. Near-term demand growth may drive increased utilization at certain plants, but increased wear and tear brings forward large maintenance investments, which in many cases will bring forward ultimate retirement dates.

At the strategic level, the delayed coal retirement strategy buys time but shifts the challenge to the future. Retirements will slow down in the near term and then accelerate again in the mid-2030s and beyond. As the large effective capacity contribution of the coal fleet rapidly retires in the 2030s, a smooth and low-cost deployment schedule for new replacement generation is essential to maintain reliability. Policymakers need to start planning and enabling investment today to ensure this future.

Gas Boom

A boom in natural gas generation is clearly underway in the U.S. power sector today. Data from the EIA **shows** nearly 30 GW of new gas generations in various stages of development will come online by 2030. A more comprehensive **survey** of development plans from S&P shows over a hundred projects totaling more than 70 GW is possible. The scale of the boom is not without precedent: Over 220 GW of capacity was deployed in the five-year period from 2001 to 2005.

Utilities and independent power producers (IPPs) are turning to gas generation to serve new demand because there is no other technology that brings as much effective capacity online, in as fast a timeline, with as much siting flexibility, under such a manageable financial profile.

Gas generation can be sited at or very near data center sites, which creates grid stability benefits and reduces overall transmission system investment costs. Meta's new 2 GW **data center** in Richland Parish, Louisiana, will host two combined cycle gas plants. Some gas generation will be deployed alongside data centers fully islanded from the grid, a model which avoids interconnection costs and delays. ExxonMobil has **announced** plans to develop 1.5 GW of fully islanded gas generation fitted with carbon capture technology and co-located with data centers, most likely sited in Texas. Siting of gas generation is somewhat constrained by the need to access pipelines for fuel. Ease of access to existing networks and easier permitting explains the strong growth in gas deployment in Texas and the Southeast. With natural gas production booming and prices at or near all-time lows, access to fuel volumes at reasonable prices is a nonissue.

The gas generation boom is creating upstream supply chain constraints. Orders for new gas turbines are rapidly piling up at major manufacturers like **GE**, **Mitsubishi**, and **Siemens**, with these firms reporting order books with delivery now stretching out past 2028. Though construction of a new gas plant can take as little as a year, with these backlogs, a project placing an equipment order today is unlikely to come online until 2030 or beyond. This order backlog inevitably includes a huge number of U.S.

projects at later stages of planning and development, so gas deployment will continue the coming years, but scaling growth rates will be a challenge.

The Solar and Storage Portfolio Play

The gas generation boom goes hand in hand with a boom in solar and storage deployment. Across different states, markets, and policy paradigms, the current economics of power generation technology favor a hybrid portfolio of gas, storage, and renewables. Gas generation delivers the effective capacity necessary to ensure demand can be served under all scenarios. Renewables, particularly solar, deliver ultra-low marginal cost electricity production on rapid deployment timelines, which improves overall portfolio costs, improves speed-to-power, and reduces the emissions profile of projects. Battery storage adds value by smoothing operations through renewable ramping periods, delivers ancillary services like frequency regulation at low cost, and brings option value that improves the overall economic and reliability profile of a generation portfolio.

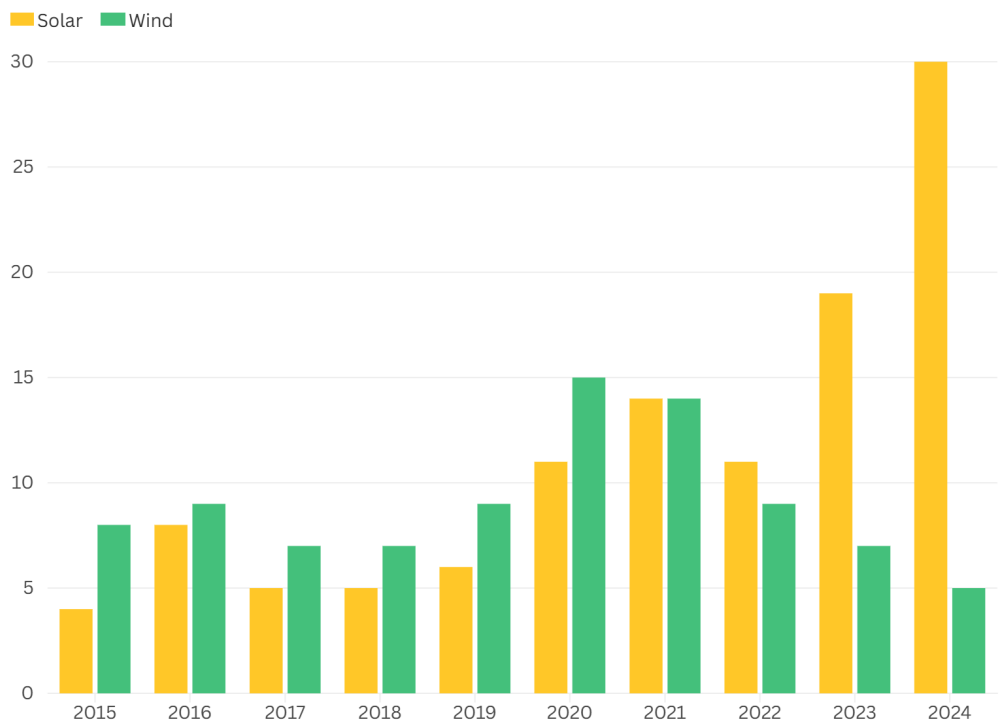
Dominion Energy, the utility that serves the Northern Virginia data center market, provides an illustrative example. Its 2024 IRP includes **plans** for 6 GW of gas, alongside 12 GW of solar, 6 GW of offshore wind, and 4.5 GW of storage by 2039. Plans from other major integrated utilities like **Georgia Power** and **Duke Energy** also display a similar portfolio approach.

Solar is rapidly coming to dominate the overall market for new generation capacity and increasingly overshadows wind's contribution. A record 30 GW of solar were **deployed** nationally in 2024; in contrast, wind deployment was at its weakest since 2014, at just 5 GW. Transmission system congestion in the nation's best wind resource regions (e.g., the Great Plains) creates long and costly interconnection processes and is a major obstacle to new wind generation. Meanwhile, solar paired with storage, directly on-site or in portfolio, has been shown to greatly **improve** the project value to electricity buyers, which has made such projects more attractive to developers and financiers relative to stand-alone wind development.

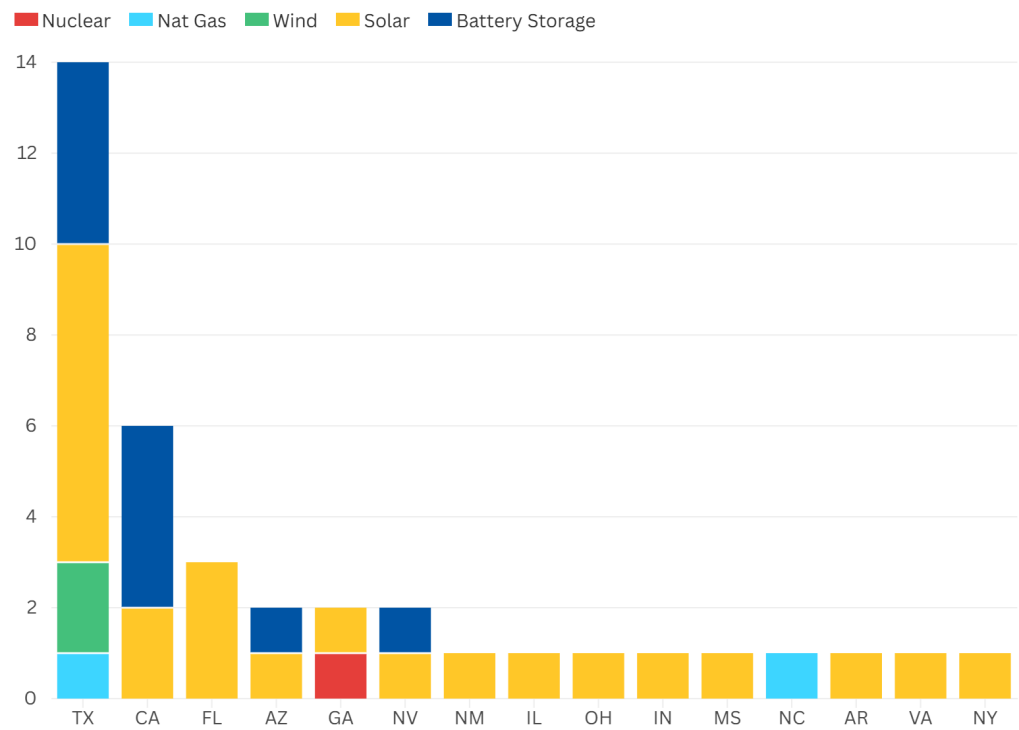
The overwhelming dominance of Texas in deploying new generation resources, with solar the dominant category, must be noted. Texas attracts investment with a low-barriers permitting environment, fast access to grid connection under the ERCOT "**connect-and-manage**" model, and plentiful land. In Texas, which is served via a competitive market rather than an integrated utility, interconnection queue data indicates incredible interest in developing solar and storage. As of January 2025, 28 GW of gas, 38 GW of wind, 153 GW of solar, and 165 GW of storage are **active** in various stages of the ERCOT interconnection queue. While many of these projects are speculative and unlikely to come to fruition, the distribution of volumes is a useful indicator: Solar and storage dominate the project development pipeline, though small behind-the-meter or fully islanded gas generation projects are excluded and would likely shift the balance slightly.

Available information about specific data centers shows that companies are building renewables to meet demand. Meta's recently announced 2 GW data center in Louisiana will be **backed** by 1.5 GW of **solar procurement** along with natural gas plants. Project Stargate, a **joint venture** between OpenAI, Oracle, and SoftBank, is anticipating data centers at the 5 GW scale. The project's first site in Abilene, Texas, will be supplied by **solar and storage projects** developed elsewhere in the ERCOT grid alongside on-site gas generation.

Figure 2: Diverging Fortunes for Wind and Solar, by Year and by State



Source: Preliminary Monthly Electric Generator Inventory, U.S. Energy Information Administration, February 2025, <https://www.eia.gov/electricity/data/eia860m/>.



Source: Preliminary Monthly Electric Generator Inventory, U.S. Energy Information Administration, February 2025, <https://www.eia.gov/electricity/data/eia860m/>.

Turning Point for Nuclear

The next five years will be dominated by deployment of large volumes of gas generation, solar, and storage. What, then, is the role of nuclear power? A series of commercial deals announced in 2024 signaled that nuclear power will also be a winner in the new era of electricity demand growth. But nuclear remains—for now—a fundamentally slow-moving technology whose primary contribution will be post-2030.

The first and now easily overlooked shift in nuclear power is the certain end to the era of premature nuclear retirements based on economics. As **recently** as 2021, over 10 GWs of reactors were planning for or at risk of early retirement. The Palisades nuclear plant was shuttered in May 2022 just months prior to the release of ChatGPT in November of 2022, which in many ways marks the start of the AI-fueled electricity demand boom.

The first pathway to “new” nuclear power is through reactor restarts. Microsoft’s **deal** with Constellation, the largest nuclear operator in the country, to restart Three Mile Island Unit 1 will bring 835 MW of high effective capacity generation to southeastern Pennsylvania in 2028. Importantly, the plant is located very close to the Northern Virginia computing cluster. The Palisades project in Michigan (800 MW) is slated to **return to service** as early as October 2025. A restart at the Duane Arnold nuclear reactor in Iowa (600 MW) is under **consideration**, but no final investment has been announced. Capacity from nuclear restarts is structurally limited however because all other retired reactors are too far along in decommissioning to be brought back online.

Upgrades at existing nuclear plants can deliver relatively small volumes of incremental new capacity. A **recent deal** between the U.S. General Services Administration (GSA) and Constellation will help finance upgrades at existing nuclear plants. In 2023, Constellation **announced** an \$800 million upgrade investment at two Illinois nuclear plants that will deliver an additional 135 MW of capacity. In total, the Nuclear Energy Institute **estimates** that upwards of 3 GW of new upgrades are possible.

Truly new nuclear projects will commence in the next five years. Several first-of-a-kind reactor projects are slated to finish by roughly 2030. These include Department of Energy (DOE)-**supported** advanced reactor designs developed by firms like **Kairos**, **X-Energy**, and **TerraPower**. These smaller-capacity and easily replicated (in theory) designs have the potential to radically alter the economics of nuclear energy from that of megaproject to something comparable to a gas-fired combined cycle. It is this theory of scaling that has attracted investment from tech firms like **Google** and **Amazon**. But policymakers should not expect perfect performance from day one from first-of-a-kind reactors. There will **inevitably** be early operational learning and design iteration periods before true commercial scaling commences. Significant contributions to the national generation mix from this segment can only be expected by the mid-2030s.

New large reactor projects, most likely utilizing the AP1000 reactor technology deployed at the recently completed Vogtle plants, are increasingly possible but not certain. New hyperscale data center clusters with demand up to 5 GW in size would appear to be natural matches for large-scale reactors. Abroad, the United Arab Emirates’ **deployment** of gigawatt-scale reactors is attracting data center investment from hyperscalers and is a model for the strategic value of large-scale nuclear in the AI era.

Despite a clear economic and strategic value proposition, the sheer size of the capital investment and cost-overrun risks loom large. The **final cost** of the recently completed Vogtle 3 and 4 reactor projects was \$32 billion, which includes \$18 billion of cost overruns. Illustrative of the challenge are recent **comments** from the CEO of Entergy, which operates multiple nuclear reactors, on prospects for new nuclear projects: “The size of the potential plant could be bigger than the entire balance sheet of the existing company, which just gives you a sense for the scale of risk that might be there for that operating company.”

Absent significant policy or commercial developments, it is not certain that new large-scale reactor projects will emerge. Restarting construction at the half-finished V.C. Summer reactor project in South Carolina is a **possibility**. Additional units at Vogtle in Georgia, reactors 5 and 6 at the plant, is likewise a plausible option. Stephen Kuczynski, former Southern Company nuclear chairman who oversaw the completion of Vogtle 3 and 4, **recently characterized** construction risk as “exaggerated” given the enormous and expensive lessons learned at the Vogtle projects. New entities like **the Nuclear Company** propose to innovate on the commercial model as an integrated project developer of large-scale. Combined with growing state-level policymaker **interest**, this indicates a plausible path forward, but more policy assistance may be needed.

Regardless, even in a best-case **construction scenario**, a new AP1000 project will take six years or more, resulting in the earliest possible contribution to the resource mix starting in the early 2030s. A steady scaling of nuclear supply chains, workforce, and technology maturation is crucial for nuclear to play a role in smoothing coal (and existing nuclear) retirements in the 2030s and beyond. Nuclear will play a limited role in the near-term speed-to-power era but could deliver enormous economic and strategic value to the nation over the medium and long term.

Federal Electricity Policy in an AI Era

What can the federal government do to ensure that the United States can power data centers and win the global race for AI? Federal policy must both address the near-term speed-to-power moment and set a long-term course toward a lasting advantage in electricity supply. In the speed-to-power era, permitting, siting, and other permissions are key areas where federal policy can help, while the generation investment choices will largely be made by the private sector and state policy makers (and rely mostly on a gas, solar, and storage expansion).

But in a sector defined by long lead times and long-lived infrastructure—a new nuclear plant or high-voltage transmission line can comfortably last 80 years—policymakers must keep an eye on the future. Investment decisions during the next several years will determine whether the U.S. grid in the 2030s and beyond allows for unconstrained electricity demand growth, at globally competitive prices, with a world-leading reliability profile—or if dramatic load growth leads to instability and internal conflict over a scarce resource.

Federal policy must also work within the framework of energy federalism. Securing U.S. dominance in AI technology is a clear national strategic priority which only federal policymakers are positioned or authorized to pursue. But federal policymakers face a jurisdictional dilemma. Electricity supply—the gating constraint on continued U.S. AI dominance—is primarily the **domain** of state-level authorities. By virtue of the long-standing **Federal Power Act**, authority over retail rates and utility generation

investments lies primarily with state policymakers. Rather than radically altering this framework, federal policy should focus on greatly improving the option set for state policymakers.

Lastly, a key area for attention is minimizing cost inflation for existing ratepayers. Electricity prices are rising rapidly, recently **outpacing inflation**. State legislatures and public utility commissions (PUCs) are facing a wave of **utility investment requirements** that translate into increasing rates. Wherever possible, federal policy should enable and encourage policy that lowers costs for generation and grid investment and reduces ratepayer exposure to investment directly tied to data centers. In cases where projects deliver clear national strategic value in the AI race, federal funding should buy down project costs to reduce ratepayer cost inflation.

Enabling the Speed-to-Power Era (2025-2030)

For the power sector, five years away is tomorrow. Demand growth over the next five years will be almost entirely served by projects already under development or construction. The data indicates clearly that generation deployment will be dominated by gas, solar, and storage. In the near-term, federal policy can primarily assist in clearing obstacles to deployment.

EMERGENCY SITING, PERMITTING, AND PLANT RETIREMENT DELAY AUTHORITIES

President Trump has already signed executive orders declaring an **energy emergency** and establishing a new **National Energy Dominance Council**. These authorities should be directed toward improving the permitting environment for generation projects that are under development in an all-of-the-above generation strategy. Fast-tracked permitting for the gas midstream, electric transmission, and electric generation projects would support speed-to-power for AI data centers. Support for enhanced geothermal on federal land is crucial for a nascent, but potentially **globally competitive**, American technology.

Most coal power plants that are operating today will likely remain open for the near term based on the new economic and reliability value proposition in the speed-to-power era, independent of the use of emergency authorities. In exceptional cases, the use of emergency authorities may be justified where state policy forces coal plant closures that raise reliability risks. Nonetheless, emergency powers are a short-term solution and should be supplemented with support for new generation that will serve the long-term multi-decadal demand growth challenge (see below).

The Northern Virginia computing cluster should be the primary focus of emergency authorities, as it is the region facing the most severe constraints on data center expansion. The administration should consider fast-tracked permitting for generation resources in the region, including offshore wind under development off the Virginia coast, which will improve speed-to-power for the strategically vital Northern Virginia computing cluster.

Emergency authorities should also target siting approval for late-stage high-voltage transmission projects that, once completed, will create room on the grid for new generation and demand resources. Focus should be paid to transmission projects that improve integration of the Northern Virginia computing cluster with new and existing generation in surrounding states. PJM has approved a series of transmission projects for this express purpose that in many cases are held up by state-siting and permitting hurdles. This authority should also consider transmission projects in the emergent demand

clusters in the Midwest, Southeast, and Southwest, which are serving a combination of strategically vital data centers, semiconductor fabrication, and battery manufacturing loads.

CO-LOCATION AND ISLANDING

The focus on speed-to-power has resulted in a strong trend toward co-locating data centers directly on-site alongside power plants. Siting new generation projects alongside new data centers is a widely pursued development strategy that poses no significant policy question. In contrast, siting new data centers alongside existing generation, as **proposed** at the Susquehanna nuclear plant in the PJM market, raises significant reliability and affordability concerns. The Federal Energy Regulatory Commission (FERC) rejected the Susquehanna proposal on narrow technical grounds but has yet to issue a formal, broadly applicable **policy** on the issue.

Numerous merchant-operated nuclear power plants in the 13-state PJM market could likely pursue similar deals if such co-location deals were okayed by the FERC. This path would radically improve speed-to-power for data centers in the mid-Atlantic market but also raise considerable reliability risks. This would in effect look like the sudden retirement of a large amount of generation from the grid without any obligation to bring on new replacement generation resources. Prices in PJM's **capacity market** would soar (if they are **allowed** to), and ratepayer prices would rise in response. The Trump administration needs to weigh the reliability and affordability risks on the one hand versus the race for AI dominance on the other.

A path forward might grant the DOE a time-limited window (e.g., through 2030) to approve co-locating at existing nuclear plants on a case-by-case basis, based on reliability assessments. One option would be to approve such arrangements only if those deals include firm plans and financial commitments to begin construction of equivalent new generation resources. Such plans could include federal support (see below). This could potentially thread the needle between speed-to-power for AI data centers and reliability.

Full physical islanding of power generation and data centers in gigawatt scale or larger “microgrids” is a way to accelerate private capital investment and improve speed-to-power. From a policy perspective this path is **attractive** because it carries no financial risk to ratepayers and poses no risk to grid reliability. Federal policy can help by clarifying that these private grids would not be subject to FERC oversight, given that they are purely commercial arrangements between private businesses. It would then lie with state policymakers to legalize such arrangements under state law and establish light-touch PUC oversight.

Building a Strategic Electricity Advantage Era (2030 and Beyond)

Federal policy in the near term can dramatically alter sectoral trajectory in 2030 and beyond for the better in terms of costs, reliability, and global strategic energy advantage. Solar will run into land-use and permitting constraints, especially outside of Texas and east of the Mississippi, where a significant volume of new data center demand is sited. It is unlikely that any state or market will be able to match the rapid interconnection rates achieved in Texas under the connect-and-manage model absent significant, slow, and politically challenging market restructuring.

Gas deployment will face delays and cost increases sourced from the turbine backlog. More importantly, in an age of liquefied natural gas exports, a growing domestic gas burn in the power sector competes with growing high-margin exports for natural gas. The fundamental **basis of energy security is in variety**, and growing reliance on a single fuel source in the power sector—in this case natural gas—eventually veers into overreliance. Though the United States possesses vast natural gas reserves, wellhead prices will eventually climb, and this will directly translate into higher electricity prices. The United States would be wise to cultivate diversity in the electricity sector, which would have the bonus value of freeing up gas volumes for high-margin overseas exports.

With these principles, constraints, and risks in mind, federal policy should focus on developing nuclear power to anchor a long-term global electricity supply advantage that supports AI dominance. Abroad, data center development is increasingly likely to flow to countries with existing or growing nuclear capacity such as China, France, Japan, and the United Arab Emirates. A nuclear-centric AI energy strategy provides the additional benefit of ensuring China does not grow to dominate the global nuclear power market as it already has with solar and storage. Over the last decade, China **built** 27 nuclear reactors compared to two in the United States, and it has another 23 reactors in various stages of construction; the United States is at risk of being left behind. Policymakers should act today to enable a post-2030 power sector that enables reliable, low-cost, demand expansion.

NUCLEAR COMPUTATION HUBS

Nuclear computation hubs would direct federal resources to states interested in both developing new nuclear power and attracting data center investment. A **10-state coalition** launched in February 2025 indicates the growing appetite for a state-led, federally supported model. States want nuclear energy but are reluctant to expose ratepayers to the risk of cost overruns. Coordination challenges hamper an alternative model that brings data center developers together around a multi-plant, multistate investment plan. Nuclear computation hubs roughly modeled after the **DOE's Hydrogen Hubs** program would cut through these hurdles. The **slow development** that has characterized clean hydrogen hubs is primarily a function of limited financial upside and investor appetite in a nascent market. In contrast, nuclear computation hubs would rapidly attract vast amounts of private capital eager to invest in the economic opportunity represented by AI and the boom in computation and electricity demand.

Nuclear computation hub applications would likely be partnerships between state energy offices, data center developer and operators, and a power developer—either an IPP or an investor-owned utility—targeting sites capable of hosting a 2 GW data center and 2 GW or more of nuclear capacity. Sites should also have plausible access to high-voltage transmission and access to additional sources of generation (e.g., gas, solar, geothermal, storage) which can support data center operations while nuclear construction proceeds.

Selected hubs would receive access to federal loan guarantees (under the DOE Loan Programs Office or equivalent authorities), grant funding for pre-Final Investment Decision site development work, federal cost sharing for high-voltage transmission investments needed to connect the cluster to the grid, expedited federal permitting, and, potentially, DOE nuclear offtake (see below). States would be encouraged to establish nuclear and data center workforce development plans for engineers, welders, and electricians, which federal funds could further support. Finally, federal support could

be made contingent on states streamlining their permitting processes for energy and infrastructure more broadly.

DOE ANCHOR OFFTAKER AUTHORITY AND NUCLEAR PROCUREMENT TARGET

As part of the Infrastructure Investment and Jobs Act, Congress created a new **anchor tenant authority** which enables the DOE to buy capacity rights (or “offtake”) in merchant transmission projects. Anchor tenancy by the federal government enables private transmission projects to secure funding from private capital markets and attract other capacity offtakers, which speeds overall deployment timelines. As other customers crowd in to contract offtake from the new transmission project, the government can surrender or auction off its contracted volumes.

Congress could authorize and fund the DOE to pursue the same model for new nuclear power. In such a model, the DOE, optionally working in consort with a federal Power Marketing Administration or the Tennessee Valley Authority, would enter into power offtake contracts with nuclear project developers. As the construction period proceeds towards commercial operation, offtake capacity can be sold off in part to private firms (hyperscalers, semiconductor fabs, etc.) or transferred to rate-regulated utilities so that the broader rate base can access the benefits of nuclear power at no risk of cost overruns. To protect taxpayers, offtake contracts should be entered into at market rates and terms. Risk sharing should be authorized insofar as it is shared across parties; this authority should not be implemented as a form of cost overrun insurance. A target of contracts supporting 10 GW of new nuclear construction underway by 2030 would radically expand the domestic nuclear construction program and ensure the 2030s are an era of rapid nuclear power growth and U.S. nuclear power leadership at home and abroad.

STRATEGIC ELECTRICITY PRODUCTION SITES ON FEDERAL LANDS

An **executive order** issued by President Biden in the closing days of his administration directed the Department of Defense, the Department of the Interior, and the DOE to identify and prepare federal sites for data center development, leveraging existing infrastructure and streamlined permitting authorities. This order should be recast with a primary focus on identifying sites for nuclear, geothermal, and solar generation. Federal land combined with fast-tracked federal permitting could be attractive for data centers only if the site delivers competitive speed-to-power. The “**three pillars of additionality**” (new clean supply, hourly matching, and deliverability) clean-energy mandate embedded in the Biden administration executive order should be scrapped to allow gas generation to be deployed as part of a portfolio power solution that prioritizes speed and flexibility. Identified sites should be made available for partnership and participation in state-led nuclear computation hubs to improve opportunities for states with large amounts of federal land.

STRATEGIC GRID INVESTMENT

The Infrastructure Investment and Jobs Act appropriated \$10.5 billion to the DOE to establish a Grid Resilience and Innovation Partnerships (GRIP) fund. Through two rounds of funding, The Grid Deployment Office has **disbursed** \$7.6 billion for 105 projects, including smart grids, renewable energy interconnection, and emergency repair projects in response to Hurricane Helene.

The program’s remaining funds should be narrowly focused on high-voltage grid investments that support the strategic goal of rapid data center interconnection. Large hyperscale computing clusters and large generation projects (e.g., combined cycle natural gas or nuclear power plants) both must

be sited close to high-voltage transmission. It is no coincidence that Meta's 2 GW Richland Parish data center in Louisiana is sited only a few miles from a branch of the U.S. Southeast's 500 kilovolt (kV) backbone transmission system. Likewise, AEP utilities in Ohio and Illinois are attracting data centers in large part due to the existing 765 kV grid system in the region.

Utilities across the country are proposing **investment** in high-voltage substations and transmission lines to support data center demand growth, and federal dollars should be deployed to reduce or eliminate ratepayer exposure to these strategically vital investments. Unlike generation, the costs of which can be easily assigned to a single large datacenter, grid investments network infrastructure whose costs and benefits are spread widely. Offsetting portions of this AI-based investment with federal dollars is key to reducing costs for ratepayers and advancing the national interest. The remaining \$2.4 billion is nowhere near sufficient to accomplish these goals. Congress should consider replenishing and expanding this fund to support the proposed nuclear computation hubs and national energy transmission corridors (see below).

National Interest Energy Transmission Corridors

Interstate energy transmission infrastructure, be it via pipeline or wire, provides broad long-term strategic benefits to the nation. Long-term policy, permitting, and political hurdles to all types of energy transmission infrastructure have undermined energy security and competitiveness.

The existing National Interest Electric Transmission Corridor (NIETC) **authority** should be expanded into a National Interest Energy Transmission Corridor authority that applies to both gas and electric transmission projects. Enabling legislative language should be streamlined to give the secretary of energy wide discernment to identify projects that serve the strategic national interest as set forth by the president. If the federal authority is invoked to site a project based on strategic national interest, then it makes sense that federal funding should likewise be deployed to pay for that national strategic value and reduce or eliminate ratepayer impact. A reformed NIETC authority would require that projects are given access to federal funding via grants (e.g., GRIP funds), low-interest loans (e.g., DOE Loan Programs Office), or anchor tenant contracts. Selected projects should also receive fast-tracked emergency permitting.

This authority could target pipelines and electric transmission projects that improve speed-to-power for existing computation clusters (e.g., Northern Virginia) and emerging computation clusters in the Midwest, Southeast, and Southwest. For example, this authority should be used to authorize and partially fund the **Piedmont Reliability Project in Maryland** which has been approved by PJM but faces political challenges at the state level. This and similar projects will boost desperately needed transmission capacity between the Northern Virginia computing cluster and the Three Mile Island nuclear plant, as well as bolstering access other firm generation resources in Pennsylvania and the Midwest. It could also be used to advance energy transmission projects which support nuclear computation hubs or to deploy pipelines which lower costs and improve reliability in pipeline constrained regions.

Conclusion

For decades, U.S. energy strategy has revolved around U.S. exposure to global oil markets. Abroad, this resulted in a focus on oil-shipping sea lanes, most notably the Persian Gulf. Domestically, this resulted in

a focus on energy independence. As an organizing principle, this is increasingly out of date. The United States has been a net energy **exporter** since 2019, and in 2024, it was the world's largest producer of both **oil** and **natural gas**.

The rise of AI has elevated electricity supply to a new level of strategic importance. A new long-term U.S. energy strategy should seek to establish global dominance in electricity supply comparable to the achieved global dominance in the oil and gas sector. The present reality of electricity scarcity that inhibits AI progress should be transformed into a long-term position of global dominance in electricity supply.

Scaling of this sort is achievable: In the decade between 1982 and 1991, U.S. electricity consumption grew by about 800 TWh, and the power sector built 43 nuclear reactors totaling 52 GW of capacity. All of this was accomplished without the aid of any modern digital engineering, manufacturing techniques, or construction technology, let alone AI itself. Whether it is nuclear, gas, solar, storage or geothermal the needed technology exists. Simply put, the engineering and technology challenges associated with meeting AI energy demand are not difficult. The onus is on policymakers to break through the status quo and unleash a future of U.S. electricity supply dominance. ■

***Cy McGeady** is a fellow in the Energy Security and Climate Change Program at the Center for Strategic and International Studies (CSIS) in Washington, D.C. **Joseph Majkut** is director of the Energy Security and Climate Change Program at CSIS. **Barath Harithas** is a senior fellow in the Economics Program and Scholl Chair in International Business at CSIS. **Karl Smith** is an economic policy consultant specializing in AI.*

The authors would like to acknowledge the crucial assistance from Bridgette Schafer and Rebecca Riess.

This report is made possible by the generous support of OpenAI.

This report is produced by the Center for Strategic and International Studies (CSIS), a private, tax-exempt institution focusing on international public policy issues. Its research is nonpartisan and nonproprietary. CSIS does not take specific policy positions. Accordingly, all views, positions, and conclusions expressed in this publication should be understood to be solely those of the author(s).

© 2025 by the Center for Strategic and International Studies. All rights reserved.