NATIONAL SECURITY, TRADE AND TECHNOLOGY

VERIFYING THE STACK

How to Promote Global Adoption of US AI While Preventing Adversary Exploitation

POLICY-IMPLEMENTATION WHITE PAPER | JULY 2026

ALVAREZ & MARSAL

LEADERSHIP. ACTION. RESULTS.℠

VERIFYING THE STACK

How to Promote Global Adoption of US AI While Preventing Adversary Exploitation

POLICY-IMPLEMENTATION WHITE PAPER | JULY 2026

Alvarez & Marsal National Security, Trade & Technology

Informed by the "Verifying the Stack" Policy-Implementation Convening[1]

Washington, D.C. | June 3, 2026

Co-organized by A&M NSTT[2]*, Akin Gump LLP, and the Institute for Security and Technology*

1 The Verifying the Stack convening brought together approximately 100 senior practitioners from across the AI policy, technology, national security, investment, human rights, and legal communities for a half-day program examining the policy imperative (WHY), the operating environment (WHAT), and the implementation architecture (HOW) for promoting global adoption of US AI while preventing adversary exploitation. The convening operated under Chatham House rules and was not recorded. This white paper does not represent the views of any individual participant, panelist, or their respective organizations, and no statement in the paper should be attributed to any participant. The paper reflects the analysis and conclusions of the authors, informed by the perspectives and operational evidence surfaced through the convening and through related discussions with practitioners across the policy, technology, operator, capital, and human rights communities. It is offered as a contribution to the shared understanding of the operating environment and the elements of a practical architecture for reconciling the dual imperatives of deployment speed and adversary denial.

2 Disclosure of interest. A&M NSTT serves as the US Department of Commerce-approved Third-Party Security Consultant for the KAUST Shaheen III GPU partition deployment cited in this paper, and, while at another professional services firm, assisted in the design and implementation of the Regulated Technology Environment underpinning the G42 Assurance Compute Framework also cited in this paper. Among other items, the paper proposes that a third-party oversight role be institutionalized as a risk-tailored structural element of the emerging AI deployment regime. Readers should weigh the authors' analysis in light of this interest. The authors seek to argue from the substance and the operational evidence. The paper's recommendations on the design of the TPSC regime, including accreditation, conflict-of-interest controls, concentration management, and liability allocation, are offered as open design questions for the broader policy and standards community, not as questions on which the authors should have privileged voice.

CONTENTS

EXECUTIVE SUMMARY

The United States is navigating a decisive point at the intersection of AI technology, infrastructure, security, and policy.

AI is the pivotal technology of our era, and a global race is underway for technological dominance, market share, and the ability to shape the rules and standards that will govern AI worldwide for a generation. Whoever wins that race will command compounding advantage in economic productivity, social organization, and strategic capability. The race is fundamentally for full-stack AI deployment: chips, servers, data centers, storage, power, cloud, models, and the trusted relationships that sustain them. It is, in other words, a competition to determine whose infrastructure becomes the foundation of the emerging AI era. The stakes of that competition are difficult to overstate.[3]

Over $1 trillion in new AI infrastructure demand is projected through 2027.[4] Large-scale AI programs across the Middle East and Indo-Pacific are translating that demand into hyperscale data center builds, national champion platforms, and multi-billion-dollar investment frameworks anchored to the American technology stack.

The policy debate in Washington has organized itself around two policy objectives that may appear at odds: **"Running Faster"**: accelerating American AI deployment to win the global technology race, and **"Capability Denial"**: preventing adversaries from exploiting US technology to accelerate their own programs or achieve decisive strategic capabilities. This paper argues that these objectives, while understandable given the real tensions involved, are complementary. Both are correct in their diagnosis, and they are not inconsistent. The United States must both run faster and deny exploitation, simultaneously, by design.

This paper is written in the context of the June 3, 2026, convening, "Verifying the Stack," which brought together pragmatists from policy, technology, and integration to address the WHY, WHAT, and HOW of this challenge. It is intended as supporting architecture for that conversation, not a single silver bullet, but a modular foundation upon which multiple complementary answers can be built.

3 This paper frames the AI build out as a competition. The authors acknowledge that other framings are possible and, in some respects, more complete. The global AI future could be conceptualized as a field that multiple nations grow together, producing mutual benefit through interoperable systems and shared standards. There are scenarios in which the optimal outcome is not zero-sum dominance but collaborative leadership. This paper is aware of and respects those alternative framings. It adopts the race framing because the decisions being made now about whose infrastructure, standards, and governance architecture become foundational are not abstract. They carry real consequences, commercial and technological, but also moral and security-related. The values embedded in the infrastructure that powers the AI era will shape the future lived experience of humanity. In the authors' view, those stakes justify the urgency and directness of the competitive framing, while remaining open to a future in which that race produces conditions for broader collaboration.

4 "Nvidia CEO Jensen Huang: $1 trillion in chip sales coming," Axios, March 16, 2026.

Two such complementary answers developed in this paper are:

Trusted Partner Coalitions. At the policy-diplomacy level, recent proposals including the Pax Silica framework and the MATCH Act reflect a growing recognition that the United States must move beyond a strictly restrictive export-control paradigm toward durable networks of allies and partners aligned across compute, cloud, semiconductor, and AI infrastructure. These coalitions will increasingly depend on verifiable assurance across the full stack: mechanisms that make trust technically measurable, auditable, and enforceable. The strategic challenge is not just to restrict adversary access to frontier capabilities but to proliferate US-led technologies across trusted partners while ensuring that those systems can be securely integrated, deployed, and governed.

Verification-by-Design Implementation. Integrating security, compliance, and trust into the full AI stack from the first deployment decision is the implementation-layer mechanism that reconciles both imperatives in practice. It enables the United States to promote deployment at the speed and scale the strategic competition requires while maintaining verified, enforceable controls against adversary exploitation.

The operationalized expression of this principle is **Verification-by-Design**: the architecture in which verification mechanisms are so deeply integrated into the deployment that the system continuously produces the evidence of its own compliance. Verification is the mechanism; assurance is the outcome it produces. This paper articulates the verification architecture that makes assurance possible at scale.

Together, Trusted Partner Coalitions and Verification-by-Design are complementary elements of a single architecture: the policy layer creates the framework, and the implementation layer makes it credible. Verification produces assurance: the durable confidence that enables regulators, allies, technology providers, and capital to extend and sustain trust.[5] Whether the bet pays off depends less on the rules written in Washington than on what happens when controlled technology meets real-world deployment conditions in Riyadh, Abu Dhabi, Singapore, and beyond.

The evidence that this works is operational. KAUST's Shaheen III GPU partition is the fastest supercomputer in the Middle East and 22nd globally on the June 2026 TOP500 list,[6] conducting cutting-edge AI research across multiple scientific domains on a verification architecture derived directly from BIS export license conditions, meeting the timelines of the OEM, BIS, and KAUST's own research plan simultaneously. G42's Assurance Compute Framework, operating under a US government-validated Regulated Technology Environment, has produced world-class AI models including JAIS 2 and Med42.[7]

These deployments demonstrate that trusted architecture does not constrain performance. It enables it. They are flagship proofs, not yet a population. Both benefited from extraordinary enabling conditions: head-of-state-level sovereign commitment, multi-year design horizons, individually negotiated license conditions, and sustained advisory capacity. The next generation of large-scale AI programs is unlikely to enjoy the same conditions. The policy task, addressed in Section V, is to translate what these flagships prove is possible into a more standardized pathway that the next ten programs can follow.

The paper proceeds in three parts. **WHY:** the policy imperative for full-stack AI export promotion as a national security priority, set against an enforcement landscape that is escalating, not retreating. **WHAT:** the operating environment of accelerating technology, four constituencies each with a distinct stake in the outcome, and the AI build-out that is the catalytic moment where those constituencies converge. **HOW:** the portfolio of complementary architectures (diplomatic and policy, and assurance-by-design) that enable the United States to run faster and maintain control simultaneously.

The central conclusion: the combination of trusted coalition architecture at the policy layer and verification-by-design at the implementation layer is not a compliance overhead. It is the mechanism through which the United States wins the AI race while ensuring that American technology cannot be exploited by adversaries. That is how both imperatives are reconciled in practice. And it is already working.

5 A&M NSTT has explored these themes in publications including: Louis Conde, Caitlin McKibben-Golub, and Randall Cook, Export Controls and Sanctions: Navigating Escalating Enforcement Risks in an Era of Global Competition, (Export Compliance Manager, Issue 60, January 2026); Randall Cook, Louis Conde, and Caitlin McKibben-Golub, AI Technology Export Enforcement: 5 Signals Companies Cannot Afford to Miss (Alvarez & Marsal, April 7, 2026); Randall Cook, Verifying the Stack: Front Line Lessons from the World's Most Demanding AI Deployments, (WorldECR, Issue 149, May 15, 2026); and "American AI Exports Program: A&M NSTT's Comments Provide Blueprint for Trusted, Efficient, High ROI US AI Exports," Alvarez & Marsal, January 16, 2026.

6 "TOP500 List – June 2026," TOP500, June 2026.

7 "Inception, Cerebras and MBZUAI Release Jais 2 – the Next Generation of the World's Leading Arabic Open-Weight LLM," Mohamed bin Zayed University of Artificial Intelligence (MBZUAI), December 9, 2025.

I. THE ARGUMENT: COMPLEMENTARY IMPERATIVES, INTEGRATED ARCHITECTURE

A. The Run-Faster Imperative

Both camps in the AI export policy debate share the same national security objective: maintain American technological leadership while preventing adversaries from exploiting US technology to close that gap. They differ on methods, not on goals. That fused premise (that both the run-faster and the capability-denial imperatives are correct in their diagnosis) is the foundation for the resolution this paper proposes. The yin and yang of the complete answer need each other.

The "run faster" emphasis is correct. The AI race will be substantially determined in the next 10 to 15 years, perhaps less, and the compression of that timeline is what is driving the fierce urgency of capital, policy energy, and technological innovation worldwide. Whoever wins commands compounding advantage. They set the terms of the next era.

The competitive logic is a virtuous cycle. Investment in superior technology and infrastructure leads to the creation of global standards and deployment at scale. Adoption and deployment generate purchase and revenue, which generates more investment, which produces more capability, which deepens the standard and the lead. This is how technology leadership compounds. It is how the United States led the internet era. It is how it must lead the AI era. The window to establish that compounding advantage is short. Speed and credibility, not multilateral consensus, define success.

There is no way to lead the AI competition without American technology companies being the most innovative, capable, and successful in the world. Effective competition requires American innovation, capital, and adaptiveness. The verification and assurance architecture this paper proposes is not designed to constrain those advantages. It is designed to provide the trusted operating environment within which they can be fully deployed at global scale: without the disruption risk of enforcement actions, the reputational cost of diversion incidents, or the political vulnerability of operating without credible assurance. The architecture enables the companies that drive American AI leadership to deploy faster, further, and with greater confidence.

The security dimension reinforces the economic one. If American full-stack AI is not the foundation of the programs being built across the Gulf, Southeast Asia, and the Indo-Pacific, non-US alternatives will fill the gap, surrendering both the commercial opportunity and the strategic influence that comes with being the infrastructure provider. A world in which the foundational AI infrastructure is built predominantly on non-US technology, standards, and data relationships would materially reduce US and allied influence over how that infrastructure is governed. The build-out is happening now.

The competitive challenge is real. Non-US AI alternatives are advancing rapidly. Huawei's 2025 revenue exceeded $127 billion, with R&D spending of $27.5 billion. In April 2026, Huawei Cloud extended its model-as-a-service offering to nine more countries, including Singapore, Thailand, Indonesia, Brazil, Mexico, Saudi Arabia, the UAE, South Africa, and Türkiye. Huawei has also deployed roughly 300 Atlas 900 A3 SuperPoDs and announced a 500,000-NPU Atlas 950 SuperCluster for release in late 2026.[8] These are real and rapidly scaling capabilities competing for the same markets that the US AI stack must win. In some cases, those alternatives draft on the very US compute investments that the export control regime is designed to control: open-source distribution of non-US models trained on American compute amplifies the competitive pressure on US providers while obscuring the underlying compute dependencies. The implication is not that the United States should retreat from full-stack export promotion. The opposite: it should accelerate trusted deployment of American technology across allied and partner markets, while building the verification architecture that makes the dependency on American compute visible and durable. The future of global AI should be built on American compute.

B. The Capability-Denial Imperative

The capability-denial emphasis is also correct, and for reasons that go beyond the imperfect effectiveness of any single control mechanism. Export controls do not need to be impenetrable to be strategically valuable. Their purpose is not to make circumvention impossible but to shape the operating environment: imposing cost, delay, and complexity that slow the pace at which controlled technology reaches unintended end uses, and preserving a window of US and allied advantage.

Export controls on advanced AI hardware are working as a retardant on adversary technology development, and the evidence that adversaries continue to attempt to circumvent them is itself proof that those controls matter. If capability denial controls were ineffective, adversaries would not spend significant resources trying to defeat them.

The scale of attempted circumvention is substantial. Epoch AI estimated in April 2026 that between 290,000 and 1.6 million H100-equivalent chips were smuggled to the Chinese mainland through the end of 2025, with a median estimate of 660,000 units,[9] roughly a third of China's total AI compute. Journalists documented eight distinct H100 smuggling networks in mid-2024, each completing transactions worth more than $100 million.[10] DeepSeek's V3 model was trained on H800 chips (a performance-constrained variant of the H100 developed specifically for compliance with US export controls),[11] and reporting suggests its R1 model may also have incorporated H100s. SemiAnalysis has estimated that DeepSeek's infrastructure includes at least 10,000 H100s that China cannot legally possess.[12]

Moreover, the scale of Chinese dependency on American compute may extend well beyond documented smuggling networks. Recent reporting and enforcement activity suggest that a significant share of advanced AI chip demand in Southeast Asian intermediary markets may be functionally driven by Chinese end-use, routed through commercial channels that are structured to obscure the ultimate end use.

8 "2025 Annual Report," Huawei, 2025; "Huawei Cloud Rolls Out MaaS in 9 More Countries," Huawei Cloud, April 10, 2026; "Groundbreaking SuperPoD Interconnect: Leading a New Paradigm for AI Infrastructure, Keynote Speech at Huawei Connect 2025," Huawei, September 18, 2025.

9 Isabel Juniewicz, Diversion and Resale: Estimating Compute Smuggling to China (Epoch AI, April 29, 2026).

10 Gregory C. Allen, DeepSeek: A Deep Dive (Center for Strategic and International Studies, April 2025); Qianer Liu, Nvidia AI Chip Smuggling to China Becomes an Industry (The Information, August 12, 2024).

11 DeepSeek-AI, DeepSeek-V3 Technical Report, (Arxiv, December 2024). Self-reported: 2,788,000 GPU-hours on Nvidia H800 chips. H800 context: Gregory C. Allen, DeepSeek: A Deep Dive, (CSIS, April 2025).

12 Dylan Patel et al., DeepSeek Debates: Chinese Leadership On Cost, True Training Cost, Closed Model Margin Impacts (SemiAnalysis, January 31, 2025); Gregory C. Allen, DeepSeek: A Deep Dive (Center for Strategic and International Studies (CSIS), April 7, 2025).

If this is accurate, it has a strategic implication that the current policy debate has not fully absorbed: the United States may possess substantially greater traction on Chinese AI development than is currently acknowledged by either US or Chinese policy leadership. That latent leverage is unrealized precisely because the current visibility architecture does not extend far enough into the distribution chain to confirm who is using controlled compute, for what purpose, and under what governance. The policy response is not necessarily to impose more restrictive controls. It is more visibility: requiring operators, intermediaries, and end users to provide verifiable data on actual use and access, so that the leverage the United States already possesses can be exercised with knowledge and intention rather than discovered after the fact through enforcement actions.[13]

Two conclusions follow from this evidence. First, the controls are working as a retardant: China is operating at a meaningful capability disadvantage relative to what it would have without them. White House AI Czar David Sacks estimated in June 2025 that China's AI sector lagged the United States by three to six months (a gap that export controls widened and sustained but did not create).[14] Second, the controls are not working as a complete barrier: sophisticated and well-resourced actors can and do circumvent them, and the scale of circumvention is growing. Both conclusions are true simultaneously. The value of a control does not depend on its being perfect. A measure that raises cost, introduces delay, and increases complexity meaningfully shapes the operating environment even when it can be circumvented at the margins. Export controls are shaping the AI technology and deployment operating environment in exactly this way, and their imperfection does not negate their effect and value.

The national security case for maintaining the American AI compute advantage operates on three reinforcing levels. First, capability: being the first to develop frontier AI capabilities, including in cyber, biosecurity, and autonomous systems, provides a window for defensive deployment before adversaries access equivalent tools. Second, volume: even when adversaries match capability levels, superior aggregate compute enables more instances, more complex tasks, and faster iteration, an advantage that persists regardless of model parity. Third, threshold: certain capability thresholds, particularly recursive self-improvement in which AI systems accelerate the development of successor AI systems, may create compounding and potentially irreversible strategic leads for the first mover. Each of these levels depends on the United States and its allies maintaining both the technology edge and the compute volume edge, which in turn depends on the integrity of the distribution, deployment, and assurance architecture.

Capability denial extends beyond hardware diversion. American technology leadership incentivizes adversary investment in circumvention, domestic alternatives, and algorithmic efficiency, much of it enabled by access to American research, deployed compute, and distillation of US-trained models. Verified deployment architecture addresses both the diversion of hardware and the ongoing risks of adversarial access once deployed, including through remote access, vicarious use, and output exploitation. It also addresses the visibility deficit that prevents policymakers from understanding who is using American compute and for what purpose; closing that deficit would significantly strengthen the US strategic position in the AI competition.

13 The evidentiary record on the scale of adversary exploitation of US AI compute is significant but incomplete. The paper draws on publicly reported enforcement actions, DOJ indictments, credible investigative journalism, and independent research (including the Epoch AI estimates cited above). Some of this evidence is anecdotal in character; the intelligence community has not publicly released a comprehensive assessment of the scope or scale of diversion. The paper accepts the validity and significance of the diversion signal without purporting to determine its precise magnitude. Several factors support this judgment: the volume and consistency of independent reporting across multiple jurisdictions and distribution channels; the scale of enforcement responses; the legislative response, including bipartisan action on the Chip Security Act and the Remote Access Security Act; and the market and regulatory response, including BIS guidance closing identified loopholes. The authors believe that a definitive empirical assessment by the intelligence community, the Government Accountability Office, and the research community would materially advance the policy discussion, and the paper's recommendations are designed to be robust whether the ultimate assessment is at the lower or higher end of current estimates.

14 David Sacks, remarks at the AWS Summit, Washington, D.C., June 10, 2025. Reported in Reuters: "Trump's AI Czar Downplays Risk AI Chip Exports Could Be Smuggled," Reuters, June 10, 2025.

C. The Resolution: Adaptive Practicality, Not Compromise

The resolution of the tension between these two imperatives is not a compromise. It is not a matter of ignoring exploitation risk to run faster, or accepting mired deployment to achieve tighter controls. Both objectives are real and both are right. The question is not which to choose, but whether they can be pursued simultaneously. The answer is yes, but effective implementation requires the right architecture.

The challenge is both more complex and more tractable than an either/or framing suggests. Running fast and maintaining control are not mutually exclusive, but making them compatible requires running fast with knowledge of who is using what and how, at scale, and building risk-appropriate mitigation controls into the deployment architecture from the beginning. Those controls must not be rigid or bureaucratic. They need to be pragmatic, proportionate to the risk, and designed to enable trusted deployment at speed rather than obstruct it. Verification should be risk-calibrated, not uniformly applied. Where deployments involve hyperscale compute, sovereign programs, or other indicators of elevated risk, the full architecture is warranted. Routine commercial AI transactions in trusted markets require lighter-touch verification. Concentrating rigor where risk concentrates is what makes the architecture sustainable at scale.

The role of government in this architecture is not solely to restrict. Restriction is necessary where adversary exploitation risk is acute, and the enforcement record makes clear that the cost of non-compliance is rising. But restriction alone will not produce the operating environment the competition requires. The government's role is equally or more to create the incentive structures, standards, and institutional architecture within which risks are appropriately identified, managed, and mitigated, and within which American innovation, capital, and adaptive capability are unleashed to solve this wickedly diverse, complex, and dynamic problem.

The wicked AI deployment problem requires a combination of instruments to effectively address. Export controls are one instrument. Standards are another. Incentive structures that reward risk-appropriate behavior, rather than simply penalizing non-compliance, are a third. Independent assurance of complex AI deployments by qualified market participants is one expression of this principle: a market-driven solution to a problem that regulation alone cannot address at the speed and scale the competition requires. A risk-adaptive architecture, appropriately calibrated to risk across the whole system, is what is needed: not one size fits all, but a coherent framework within which restriction, incentive, and innovation work together rather than at cross-purposes.

Verification-by-design is the principle that makes this possible. It integrates security, compliance, and trust into the AI stack from the initial deployment decision, not as a post-hoc compliance overlay but as an operating property of the system. The architecture makes compliance intrinsic rather than external. And the operational evidence from the most advanced authorized deployments in the world (KAUST and G42) demonstrates that this architecture enables deployment speed and resilience rather than constraining it. What those deployments demonstrate is not merely that verification works, but that it produces assurance: the operating condition in which compliance confidence is continuous rather than episodic. The June 2026 government-directed suspension of a frontier AI model deployed through cloud infrastructure underscores the distinction: cloud-contingent access is structurally fragile; assured, sovereign, verified compute is not.

Treating export compliance as a friction cost that slows deployment is precisely wrong. In this environment, the organizations that operationalize security, resilience, and trust most effectively are the ones that deploy fastest in steady state. Trusted architecture is not the enemy of performance; it is a necessary condition for it.

D. The Fabric of Trust: An American Competitive Advantage

There is a dimension of the resolution that deserves explicit attention because it is both underappreciated in the policy debate and central to why the American approach is structurally superior to the alternative: the fabric of trust that is intrinsic to the American model at two complementary levels.

At the policy and alliance level, the United States is constituted as a nation with roots in all nations, a global power whose strategic relationships are built on shared equities, accountability to shared human values, predictability, and a track record of honoring commitments. These are persistent features of American strategic culture that have powered coalition-building across eras and remain unique strategic assets in the AI competition. China can deploy capital and technology, but it cannot replicate the network of relationships, legal frameworks, and shared values that underpin American alliances. The Pax Silica framework, the AAEP consortium pathway, and the multilateral governance structures being built around AI deployment are expressions of this alliance architecture, accessible only to countries that operate within the American orbit.

At the implementation and verification level, the American model rests on the rule of law, predictability, consistency, accountability, and verification. These are not just governance values. They are competitive advantages. When AI operators build on American technology within an assured compliance architecture, they are purchasing more than compute. They are purchasing membership in a governance ecosystem that is auditable, transparent, enforceable, and durable. Those integral features of the American approach to capitalism, property, and innovation have differentiated and sustained America's vitality and dynamic success in the past. Those features of trust and partnered integrity are precisely what AI operators with alternatives are choosing today, and it is why KAUST and G42 have built their programs on American technology and compliance frameworks rather than alternatives.

The values dimension of this competition is inseparable from the architectural one. AI is a uniquely powerful technology, with enormous potential both to enable and suppress human dignity. The governance values embedded in the infrastructure on which AI is built will shape humanity's future. Building on trusted architecture grounded in the rule of law, transparency, and accountability is therefore not only a commercial or security consideration but a matter of the values that the technology will carry into the world. The verification and assurance architecture this paper proposes is the mechanism through which those governance values, transparency, accountability, independent oversight, and respect for individual rights, are embedded in the technology stack itself rather than asserted in policy documents that the technology does not enforce.

It is worth noting that we are in a disruptive era across multiple domains. The structures of the post-Cold War international order are being reexamined in the context of political, social, and technological change. Rigid and friction-permeated institutions will not survive this disruption. The AI Diffusion Rule is not being enforced. Multilateral export control structures including Wassenaar face new pressures. Some of what existed before will be cleared away. Yet within this disruption, there remains a through line: the persistent features of American comparative advantage that have powered American leadership in past eras can and must power its continued success in this one. Human innovation, the rule of law, the capacity to build trusted partnerships, and the integrity of commitments are not products of any single party or administration. They are durable features of American strategic culture that need to be sustained and leveraged in the AI deployment race. Verification-by-design architecture is built on those features. So are the allied coalitions that the diplomatic HOW is designed to construct and sustain. The disruption of the present moment creates the opportunity to build something better on that enduring foundation.

II. WHY: THE POLICY IMPERATIVE

US AI export policy has converged on a single strategic objective: promote global adoption of full-stack American AI while preventing adversarial exploitation.

The two halves are not in tension when the architecture is right; they are mutually reinforcing. American AI embedded in sovereign infrastructure creates technical dependencies, operational relationships, governance norms, and strategic influence that passive restriction cannot achieve. The AI competition is fundamentally a placement competition, not purely a denial competition: the United States needs its AI in the infrastructure that will shape global digital governance for a generation. The question is not whether to place it, but how to place it in ways that sustain American advantage, preserve the integrity of the technology, and build the trust relationships that prevent adversary exploitation, at scale and at speed. Assured deployment architecture is what makes that embedding trustworthy and sustainable.

A. The Policy Trajectory: Constant Equities, Adaptive Instruments

The two equities described above are constant poles of US AI export policy. Any regulatory or implementation architecture must be calibrated and adaptive enough to account for both simultaneously. An architecture that does not will not survive. The Trump administration concluded that the Biden-era AI Diffusion Rule was not well-calibrated to address those equities in an adaptive and pragmatic way: it was viewed as too bureaucratic, too friction-oriented, and would ultimately stifle the at-speed deployment that the run-faster imperative requires.

In the absence of a successor regulatory structure, the policy apparatus and the stakeholders around it are using the instruments available to drive these objectives forward. Those instruments operate on two tracks that, taken together, are fully consistent with what the objectives require.

Run-Faster Instruments. Export promotion and industrial policy operationalize the run-faster equity directly. Individual licensing decisions, government-to-government frameworks, and consortium-based pathways such as the American AI Exports Program (AAEP) are each designed to move advanced US AI into trusted partner deployments at speed and scale. The publicly reported major AI licenses granted to KAUST and G42, and the government-to-government frameworks enabling similar programs, are each a policy signal that deployment is the objective. The H200 licensing policy toward the China market reflects the same impulse. These instruments are primarily focused on enabling deployment, but each carries within it the expectation of assurance architecture as the condition of access.

Prevent-Exploitation Instruments. Legislation and enforcement are driving home the prevent-adversary-exploitation imperative in the absence of regulatory coordination. The Chip Security Act and the Remote Access Security Act establish hardware-level verification and access control standards with bipartisan support. Aggressive BIS and DOJ enforcement, escalating penalties, and a substantial requested budget increase signal that compliance architecture is the price of market access, not an optional enhancement. These instruments are primarily focused on the protection equity, but they too carry the deployment objective: the enforcement posture is designed to make compliant deployment accountable, not to stop deployment.

Taken together, the two tracks reflect the persistent, fused through lines of US AI policy: accelerated deployment and high assurance expectations are both required, advanced through different instruments calibrated to each objective while carrying elements of the other. This is the successor regime taking shape, not through a single coordinating rule, but through the accumulating weight of instruments responsive to persistent, necessary objectives.

B. The Signals Landscape

The policy and enforcement environment is converging on a clear set of expectations about how advanced AI compute can be deployed, distributed, and accessed. The signals are operational, not theoretical, and they point in a consistent direction: visibility into end use, accountability across the supply chain, and access control across all modes of delivery are becoming threshold requirements for sustained market participation. Operators who are not building to meet these expectations face escalating disruption risk: enforcement, license revocation, securities liability, and provider-level interruption that the operator cannot control. The signals below sort into three reinforcing themes.

First, enforcement is reaching further into the supply chain and higher into the corporate hierarchy. The March 2026 indictment charging individuals associated with a major server manufacturer alleged a coordinated scheme to route billions of dollars in AI servers containing controlled GPUs through Southeast Asian intermediaries before delivery to Chinese end users: assembly in the United States, transshipment through allied jurisdictions, repackaging for ultimate delivery, all of which operated undetected for an extended period.[15] The $252 million BIS penalty imposed on an advanced materials manufacturer in February 2026 for supply-chain-related violations marked the second-largest civil penalty in BIS history and established that civil exposure now scales with the magnitude of the underlying transactions.[16] BIS's substantial enforcement request for resourcing, paired with parallel DOJ activity, signals that the cost of non-compliance is rising and that the enforcement apparatus is expanding to match the scale of AI infrastructure deployment.[17]

Together, these cases demonstrate that export control risk is no longer confined to operational compliance functions: criminal liability can attach to senior leadership, and parallel securities litigation following enforcement actions indicates that export control failures are now a market-facing risk with direct implications for valuation, disclosure, and the personal exposure of officers and directors.

Second, the regime is redefining what "control" means. Increasingly, control is defined by access and use, not by the physical location of the hardware. Recent enforcement actions and pending legislation treat remote access, API-based interaction, and other indirect use as functionally equivalent to a physical export. The Remote Access Security Act would codify this principle in statute. Investigations by US and Singaporean authorities into a major Southeast Asian AI cloud provider for potentially enabling re-export of advanced chips to the Chinese mainland through regional subsidiaries and cloud-based access models illustrate how intermediary jurisdictions and commercial distribution models can obscure ultimate end use in ways that current enforcement mechanisms are not fully equipped to detect in real time.[18] The Chip Security Act's legislative sponsor has articulated a forward-looking vision in which controlled chips are subject to hardware-enforced, time-limited cryptographic licensing, analogous to the hardware-bound licensing models that govern EDA design tools and the secure boot architecture in consumer devices, verifiable, renewable, and revocable at the hardware layer independently of the governance architecture above it. While more technically ambitious than the current legislative text, this vision indicates the direction in which the enforcement architecture is intended to evolve: from controlling the point of export to controlling the conditions of ongoing use.

Third, and newest, cloud-based access to controlled AI carries disruption risk independent of operator compliance. On June 12, 2026, a US export control directive required a major American frontier AI provider to suspend access to two of its most capable models for any foreign national, including its own foreign national employees.[19]

15 US Department of Justice, indictment unsealed March 19, 2026, United States v. Liaw et al., S.D.N.Y. Reported in: "Co-founder of Tech Company Charged with Diverting $2.5 Billion in Nvidia AI Chips to China in Violation of Export Laws," CNN Politics, March 19, 2026.

16 "Applied Materials to Pay $252 Million Penalty to BIS for Illegally Exporting Semiconductor Manufacturing Equipment," Bureau of Industry and Security, US Department of Commerce, February 11, 2026.

17 "David Peters, Assistant Secretary for Export Enforcement, Bureau of Industry and Security, Testimony before the House Committee on Foreign Affairs, South and Central Asia Subcommittee," Bureau of Industry and Security, US Department of Commerce, February 24, 2026.

18 "Singapore police probe Nvidia customer Megaspeed over alleged China export violations," CNBC, October 10, 2025.

19 "Anthropic Suspends New AI Models After Government Directive" NBC News, June 12, 2026.

Because the provider could not reliably identify the nationality of every API caller, it disabled both models for its entire global customer base the same day, with no transition period. This was the first time the US government used export control authority to take a publicly deployed frontier AI model offline.

The episode is now in a contested phase: access has begun to be restored on conditioned, vetted-user terms, and the directive is in litigation. The structural lesson does not depend on how the case resolves: access to frontier capability delivered through commercial cloud infrastructure can be conditioned, narrowed, or withdrawn by government action directed at the provider, independent of any end user's compliance posture.

That trajectory is reinforced by the parallel emergence of government-managed release, in which frontier models are increasingly made available through vetted-partner or pre-release-review channels rather than open deployment: a pattern the AI Innovation and National Security Executive Order of June 2, 2026 anticipates in its framework for providing advanced models to the executive branch for evaluation. For operators building large-scale AI programs on assured, on-premises, verifiably governed infrastructure, the contrast is instructive: computational capability resident within a verified security boundary under institutional control is not subject to the same external, provider-level revocation.

Taken together, these three themes describe an operating environment in which market access is being conditioned on verifiable end-use compliance, the perimeter of liability is expanding outward into the supply chain and upward into corporate leadership, and the durability of cloud-delivered access cannot be assumed. The operators who disregard or treat these signals as a transitional inconvenience will be vulnerable. The operators who build deployment architecture addressing the expectations these signals telegraph: access verification, supply-chain provenance, end-use visibility, and institutional control of the infrastructure itself, will be positioned to deploy at speed, resilience, and scale when the formal regime catches up to the practice that is already emerging.

C. The Successor Regime Is Taking Shape

The authors of this paper believe that a formal successor rule contemplating these equities, adaptive, outcomes-based, and modular enough to apply verification-by-design principles across diverse deployment contexts, would be the most efficient and durable way to embed this architecture. However, the argument does not depend on that rule materializing. The instruments described above are creating the successor regime now. The entities that build compliant architecture now will be positioned to lead. Those that do not will face efficiency, assurance, and retrofit problems in a market where the standard is set by the most demanding implementations already operating.

It is uncertain whether a formal successor rule to the AI Diffusion Rule will be issued, or when. The anticipated successor rule was submitted for OIRA interagency review in early 2026, but withdrawn in March 2026, reportedly due to unresolved policy disagreements regarding the scope and structure of the qualification pathway, ally treatment, and the enforcement architecture. The withdrawal does not mean the successor rule is defunct. It means the interagency process may need help to chart a workable path through admittedly complex terrain.

In the interim, deployment continues and the composite direction is clear. The AAEP Call for Proposals opened April 1, 2026, and closed June 30, 2026;[20] whether the AAEP becomes the durable channel or is succeeded by a more flexible framework, the qualification expectations being articulated through this process inform the standards that large-scale AI deployments will need to meet. The Chip Security Act's passage through committee reflects bipartisan consensus that hardware-level verification must be a threshold requirement. BIS enforcement activity reflects a regulatory posture in which access to advanced compute is conditioned on compliance with the governance architecture.

Section V develops the design principles a successor regime must embed, whether codified in a formal rule or accumulated progressively through the four instruments above. The convening on June 3 and this paper are designed to accelerate that process across all four channels.

20 American AI Exports Program, US Department of Commerce. Call for Proposals: April 1, 2026, through June 30, 2026. See: "American AI Exports Program," US Department of Commerce.

III. WHAT: THE OPERATING ENVIRONMENT

The AI competition is being run in an operating environment defined by over $1 trillion in projected new infrastructure demand through 2027,[21] four constituencies with distinct stakes in how the United States competes, and a Gulf and Indo-Pacific build-out that is the catalytic moment at which those constituencies converge.

A. Four Constituencies, Converging Stakes

The operating environment defined by this technological moment has distinct implications for four constituencies, each of which has a different stake in how the United States reconciles the imperatives of deployment speed and adversary denial. Understanding their interactions is essential to designing policy and deployment frameworks that work in practice rather than in theory.

Policy and Hill. For congressional staff and agency officials, the operating environment defines the enforcement gap: the distance between what the policy requires and what the market delivers without architectural enforcement. The successive failure of regulatory frameworks that lacked implementation architecture (from the AI Diffusion Rule to the proliferating enforcement actions that followed its rescission) has created both urgency and opportunity. The successor rule, the Chip Security Act, and the AAEP qualification pathway are the mechanisms available. The question is whether those mechanisms will be designed to close the enforcement gap or to paper over it. Policymakers in this constituency need operational evidence, not theory, about what verified deployment requires and what it produces.

Technology and Industry. For AI infrastructure providers, hyperscalers, OEMs, and semiconductor companies, the operating environment defines the compliance architecture they must build into their products and their customers' deployments. The enforcement record makes clear that the compliance question is no longer optional; the only question is whether it is addressed architecturally from the start or remediated at significant cost after an enforcement action. Verification-by-design is increasingly a competitive differentiator: operators who can demonstrate trusted deployment architecture attract large-scale clients who require it, access licensing pathways that require it, and generate regulatory goodwill that protects them when the landscape shifts. The visibility imperative extends to the technology supply chain itself: as enforcement actions and reporting reveal that controlled technology may be reaching adversary end users through intermediary channels on a larger scale than previously assessed, technology providers face direct interest in verification architecture that can demonstrate the integrity of their distribution networks. The reputational, legal, and securities-liability consequences of being associated, even unknowingly, with large-scale diversion are severe.

21 Jensen Huang, Keynote Address, NVIDIA GTC 2026, San Jose, CA, March 16, 2026. Reported in: "Nvidia CEO Jensen Huang: $1 trillion in chip sales coming," Axios, March 16, 2026.

Data Center and Infrastructure Operators. For large-scale AI operators (KAUST, G42, and their equivalents across the Gulf and Indo-Pacific) the operating environment defines the compliance architecture they must demonstrate to access advanced US-licensed compute and to renew that access over time. The most sophisticated operators have recognized that compliance architecture is not a constraint on their programs but a prerequisite for the access they need. KAUST and G42 have built programs that demonstrate this at the highest level of performance. Their experience, and the architecture that makes it possible, constitutes the model the next generation of large-scale AI programs must follow.

Capital and Consortium Participants. For sovereign wealth funds, infrastructure investors, and deal teams underwriting large-scale AI platform investments, the operating environment defines the compliance and execution risk picture they must price into their deployment decisions. Operators with credible assured compliance architecture can deploy faster, attract better licensing terms, and sustain deployment at scale without the disruption risk of enforcement action or license revocation. Operators without it face a hidden risk premium that compounds over time. The capital community ultimately underwrites which operators win. That means the capital community has a direct stake in understanding what assured deployment architecture looks like and what it costs to get right from the beginning.

These four constituencies interact in ways that are mutually reinforcing when the architecture is right. Policy frameworks that establish clear qualification standards create the incentive for operators to build assured architecture. Operators with assured architecture attract capital at better terms. Capital deployed into assured programs generates the returns that justify continued investment. Technology providers that build verification-by-design into their products at the infrastructure layer enable the whole ecosystem to operate more efficiently. When all four constituencies are aligned around the same standards and incentives, the virtuous cycle of investment, deployment, and reinvestment that drives American AI leadership accelerates.

B. The Gulf Build-Out: Where the Four Constituencies Converge

The Gulf region (principally the UAE and Saudi Arabia) is the decisive inflection point where all four constituencies converge. The combination of sovereign wealth, strategic ambition, political will, and openness to trusted partnership that will power the engines of a new industrial age has created a concentration of demand, capital, and deployment opportunity that is unmatched anywhere in the world. This is not an accident. It is the consequence of deliberate national strategies (Saudi Vision 2030, UAE AI Strategy 2031) that have positioned Gulf nations as the buyers, trusted partners, and operators of the large-scale AI programs that will define regional and global AI infrastructure for years.

KAUST's Shaheen III GPU partition is the fastest supercomputer in the Middle East and ranks 22nd globally on the June 2026 TOP500 list.[22] It supports cutting-edge research across materials science, genomics, climate modeling, and large-scale AI, under what may be the most demanding compliance framework ever applied to an AI deployment. The compliance architecture was built on verification-by-design principles and met the timelines of the OEM, BIS, and KAUST's own research plan simultaneously.

Critically, KAUST built its assurance architecture by adapting existing research computing governance, including a multi-disciplinary evaluation committee that has operated since 2009, to incorporate the AI model controls, end-use screening, and data governance the export license required. This demonstrates that assurance-by-design does not require greenfield construction; it can be built on existing institutional foundations, lowering the barrier for the next generation of operators. Early research output on the platform spans wave modeling, materials science, climate modeling, genomic research, and Arabic-language AI, validating the proposition that trusted architecture enables performance rather than constraining it.

22 "TOP500 List – June 2026," TOP500, June 2026.

G42, the Abu Dhabi-based national AI champion, operates a globally distributed constellation of AI supercomputers including the Core42 Maximus cluster, an AMD Instinct MI300X system at its Buffalo, New York facility, which ranked 20th globally on the November 2025 TOP500 list.[23] Operating under the G42 Assurance Compute Framework, a regulated technology environment developed with the US and UAE governments and referenced by BIS as a model for future export license conditions, G42 has deployed world-class AI capabilities including JAIS 2, the leading Arabic large language model, and Med42, a clinical LLM that surpassed Med-PaLM on benchmark performance.[24] The framework is designed as a scalable architecture extensible to additional deployments and Pax Silica ecosystem partners: the commercial expression of the proposition that trusted deployment and global AI competitiveness are not mutually exclusive.

Additional large-scale AI programs, including G42 and Saudi Arabia's Humain, have received US government approvals for advanced Blackwell-generation compute under government-to-government frameworks. These programs represent the next wave of large-scale AI build-out that will need trusted deployment architecture from inception.

Multiple channels, individual export licensing, bilateral government-to-government frameworks, and the AAEP qualification pathway among them, are available to extend large-scale AI access on terms that protect both the operator and the US technology base. The mix of channels will likely continue to evolve as the policy environment matures.

C. Regulatory Interregnum: Risk and Opportunity

The interregnum between regulatory frameworks is simultaneously the highest-risk and highest-opportunity moment for the operating environment. The enforcement record is escalating. The legislative trajectory is clear. Multiple deployment channels are active. The successor rule is in development. The operators, investors, and policymakers who engage seriously with the compliance and assurance architecture now (before the standard is formally codified) will be best positioned when it is codified. Those who wait will face an enforcement risk and a retrofit problem in a market where the standard is set by the most demanding implementations already in operation.

23 "Core42 "Maximus-384" Cluster Secures Top 20 Ranking on the Global TOP500 Supercomputers List," HPCwire/Business Wire, November 18, 2025.

24 "M42 Announces New Clinical LLM to Transform the Future of AI in Healthcare," M42, October 12, 2023.

IV. HOW: THE INTEGRATED ARCHITECTURE

Achieving both objectives simultaneously, running faster and denying adversary capability, requires adaptive architecture. The HOW is not singular.

It is a symphony of approaches that, taken together, enable the United States to achieve both objectives simultaneously at scale. No single instrument or approach, however well-designed, is sufficient. The challenge requires pragmatists from policy, technology, and integration working in parallel on approaches that reinforce rather than substitute for each other.

This paper identifies two complementary layers of the integrated architecture: a policy layer of trusted partner coalitions, and an implementation layer of verification-by-design. Each will be developed in greater depth through the work the June 3 convening and subsequent thought leadership are designed to advance. A full resolution of the strategic challenge will require both, and potentially additional 'how' approaches the field is only beginning to define.

A. The Policy Layer: Trusted Partner Coalitions

The United States and its allies are entering a period in which AI governance, compute infrastructure, semiconductor access, and digital standards will increasingly function as instruments of geopolitical alignment. As frontier AI capabilities diffuse across jurisdictions, competition over AI is expanding beyond model development itself and into the broader systems that enable capability generation and deployment: energy, cloud infrastructure, advanced semiconductors, data governance, financing mechanisms, and the security of globally distributed supply chains.

Currently, no single country, including the United States, can independently shape the global operating environment for advanced AI systems. Durable advantage will instead depend on whether the US and its allies and partners can establish a coherent governance and infrastructure ecosystem to shape how frontier AI capabilities and the supply chains that underpin them are developed, accessed, secured, and deployed.

A coalition-based approach serves two parallel objectives: strengthening the collective capacity of trusted states to accelerate AI development within secure, interoperable, and politically-aligned environments; and reducing the ability of adversarial actors to exploit gaps among allied regulatory systems, procurement standards, export control regimes, supply chains, and infrastructure networks. Existing initiatives, including export control coordination, plurilateral technology dialogues, and emerging sovereign AI partnerships, address discrete components but do not yet constitute an integrated architecture. Frameworks such as the MATCH Act, AAEP, EDAG financing structures, and Pax Silica point toward the early foundations of one, though significant gaps remain in implementation and institutional durability.

A central policy challenge will be balancing openness with security across different categories of partners. Indo-Pacific allies are primarily concerned with supply chain resilience, economic prosperity, regional deterrence, and in some cases reducing dependence on Chinese technology ecosystems. Gulf partners are pursuing rapid AI-enabled economic diversification and large-scale compute buildouts while navigating complex relationships with both the United States and China. European partners remain focused on regulatory governance, digital sovereignty, and standards-setting authority. These differing strategic priorities imply that a single governance model is unlikely to suffice across all coalition members.

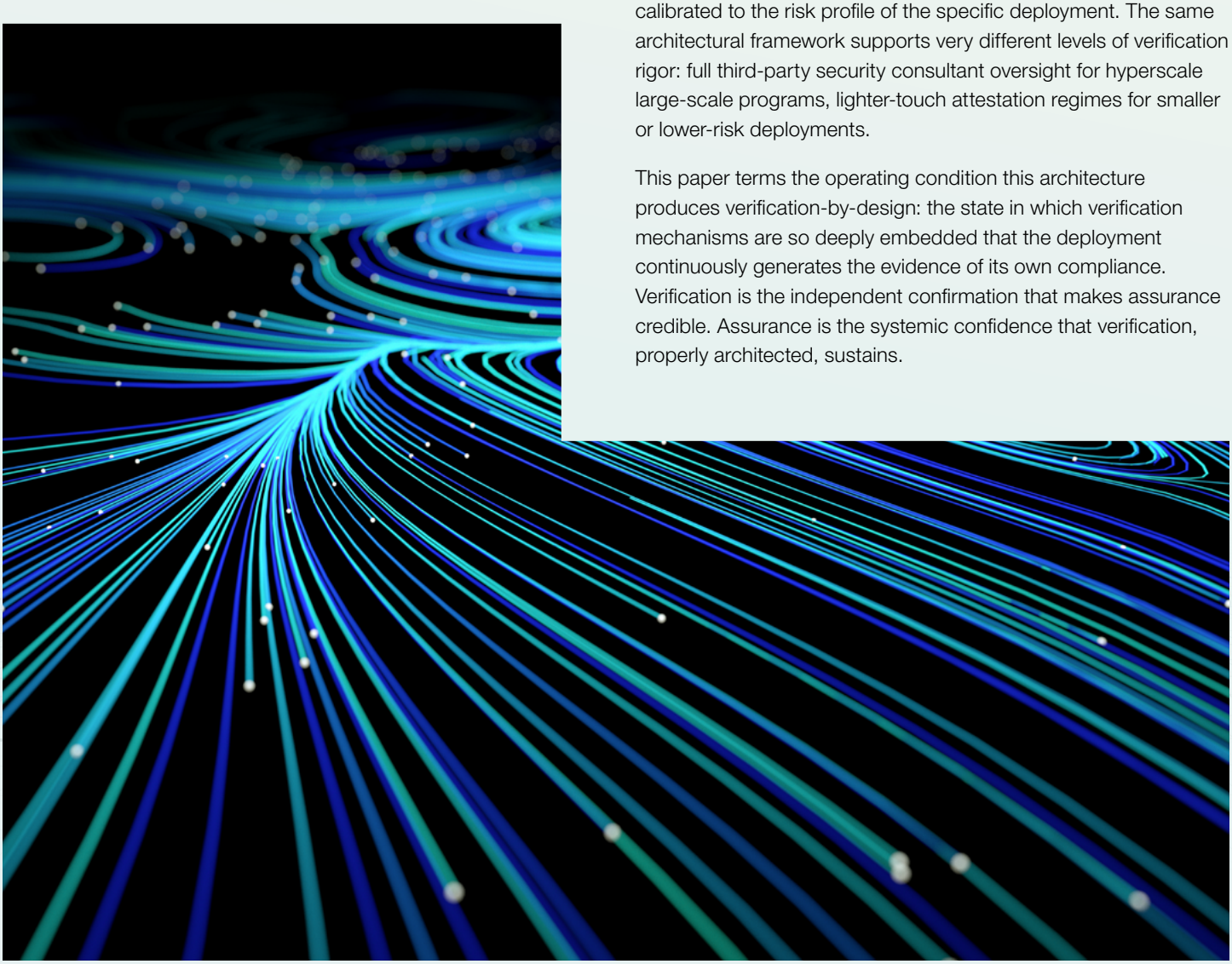
As a result, the most durable strategic and diplomatic architecture is likely to emerge through layered and partially overlapping structures. Bilateral arrangements will remain critical for sensitive technology sharing, compute access, and intelligence cooperation.

Plurilateral mechanisms will likely prove more effective for coordinating export controls, infrastructure security standards, and cloud governance practices among technologically aligned partners. Broader multilateral forums may continue to play a role in norm development, but are less likely to produce binding commitments at the speed required by frontier AI competition. Within this framework, the US retains several key sources of leverage: access to advanced semiconductors, cloud infrastructure, investment capital, and financing mechanisms provide Washington with substantial influence over allied AI development pathways. Incentives could include preferential licensing arrangements, participation in trusted compute consortia, and access to shared evaluation and testing frameworks: tools increasingly deployed alongside coordinated export controls, outbound investment screening, and verified controls limiting compute access for actors operating outside trusted networks.

B. The Implementation Layer: Verification-by-Design

Ronald Reagan's arms control maxim had a three-word core: "trust but verify." The analogy is imperfect, since bilateral arms control is different from conditions on AI security, use, and access. Reagan's discipline still translates: trust is not extended without the means to verify, and verification is not credible without architecture designed to support it. Verification-by-design is what that doctrine looks like in the AI era. It integrates security, compliance, and trust controls into the AI deployment architecture from the first design decision, not retrofitted after deployment, not triggered only by incidents, and not satisfied by periodic audits alone. The architecture makes compliance an operating property of the system rather than an external overlay. The result is a compliance posture that is auditable, continuous, transparent to the licensing authority, and renewable as the deployment scales. The depth and intensity with which each dimension is operationalized should be calibrated to the risk profile of the specific deployment. The same architectural framework supports very different levels of verification rigor: full third-party security consultant oversight for hyperscale large-scale programs, lighter-touch attestation regimes for smaller or lower-risk deployments.

This paper terms the operating condition this architecture produces verification-by-design: the state in which verification mechanisms are so deeply embedded that the deployment continuously generates the evidence of its own compliance. Verification is the independent confirmation that makes assurance credible. Assurance is the systemic confidence that verification, properly architected, sustains.



The assurance model this architecture implements is analogous to practices well established in the capital, insurance, and transaction risk management communities: independent, risk-calibrated assessment by qualified professionals that provides confidence to all parties (the operator, the licensing authority, and the capital community) that the deployment meets the required standard. The government sets the outcomes. The market provides the assurance. This is how complex, high-stakes systems are governed in mature industries, and it is the model the AI deployment regime should adopt.

The architecture addresses four risk dimensions that any large-scale AI deployment operating under US export licenses must manage:

Access Risk. Specifies who can access the system, by what means, and under what authorization, including remote, API-based, and programmatic access treated with the same rigor as physical access. Identity and affiliation screening embedded upstream of system access, with project-specific authorization rather than blanket user access.

Diversion Risk. Ensures supply chain integrity and liability across the full hardware and software stack, from procurement through deployment, including chain-of-custody documentation, periodic hardware inventory verification, and controls on the export of trained model weights.

Use and Output Risk. Defines what the system can be used for, by whom, and what it may produce, including workload-level authorization, prohibited end-use screening, vicarious use prevention, trained-model protections, and governance of agentic AI systems that can autonomously pursue complex objectives. Emerging evidence that the scale of adversary access to American compute through intermediary channels may be substantially larger than previously assessed underscores the importance of use and output governance that extends beyond the licensed deployment to the broader ecosystem of access and distribution. AI-assisted diligence tools, including automated screening of users, affiliations, and research outputs, are becoming operationally essential at the scale of current deployments.

Verification Risk. Requires continuous, independently verified compliance, not point-in-time audits, including persistent monitoring, mandatory escalation, government-validated reporting, and a trusted third-party verification authority with the mandate, access, and expertise to verify compliance continuously and report transparently.

Verification-by-Design: A Digest of the Implementation Architecture

These four risk dimensions are operationalized through six implementation pillars. The full architecture will be detailed in a forthcoming annex to this white paper.

Pillar One: Data Center and Infrastructure Security. Physical and logical controls governing the facility, hardware, and network environment. Includes biometric access, network segmentation, and alignment with the relevant US standards baselines, principally NIST CSF 2.0 and NIST AI RMF 1.0 (with the AI 600-1 Generative AI Profile), supported by the appropriate NIST SP 800-53 r5 control families for federal-grade information systems. The Chip Security Act addresses a subset; the full pillar extends to environmental controls and facility governance.

Pillar Two: Customer Access, Identity, and Use Controls. Authorization and identity verification upstream of system access, with project-specific containment. Operationalizes both Access and Use/Output risk dimensions through formal proposal, review, authorization, and monitoring.

Pillar Three: Personnel Security and Insider Threat. Screening and monitoring of all individuals with access: identity checks, background screening, need-to-know restrictions, behavioral monitoring.

Pillar Four: Ownership, Investment, and Foreign Influence. Beneficial ownership transparency, foreign investment screening, controls on governance changes. Addresses diversion risk that hardware controls alone cannot reach.

Pillar Five: Supply Chain and Asset Life Cycle Management. Chain-of-custody from procurement through decommissioning. Serial-number-level inventory, trained-model weight controls. Recent enforcement actions illustrate the severe consequences of supply chain and asset life cycle failures at this layer. The software-supply-chain dimension is equally load-bearing: practice anchors such as SLSA, in-toto attestation, SBOM formats (CycloneDX, SPDX), and NIST SP 800-218 (the Secure Software Development Framework, with its AI-specific 800-218A profile) define the artifacts a 2026-grade verification regime should expect operators to produce.[25]

Pillar Six: Monitoring, Reporting, and Certification. Continuous monitoring, anomaly detection, mandatory escalation, government-validated reporting. Independent verification and assurance providers, a role currently instantiated through the US Department of Commerce-approved third-party security consultant designation, but conceived here as a broader market function, are embedded from design through operations and are the architectural innovation distinguishing assurance-by-design from conventional compliance.

These six pillars are integrated through a unified governance architecture. Technology-enabled verification (hardware attestation, confidential computing, workload telemetry) strengthens each pillar. But technology alone is necessary but not sufficient. The governance architecture linking human accountability, independent oversight, and transparent reporting is what makes the system credible.

The organizational architecture of the operator also matters. Systems tend to reflect the structures of the organizations that build them. If compliance is siloed in a legal function separated from engineering, controls are bolted on rather than built in, and friction is maximized rather than managed.

The most effective implementations push security and compliance decision-making to the operational edge, where engineering choices are made, and embed verification into the development and deployment life cycle through reproducible, codified processes.

KAUST and G42 represent complementary models: KAUST operates a governance-dense, institution-embedded architecture; G42 operates a platform-centric, automation-driven architecture. A scalable global standard will draw from both.

25 Provenance is the same problem at both ends of the supply chain. The integrity of controlled hardware reaching trusted destinations and the integrity of the upstream inputs that produce that hardware are structurally identical challenges: both require auditable chain of custody, traceable sourcing, and a workable evidentiary standard that operates at scale. Rebuttable-presumption frameworks already in use in adjacent areas of US trade enforcement offer an instructive design parallel. They establish a default compliance expectation that shifts the burden of proof to the party seeking to demonstrate conformity, creating structural incentives for transparency and documentation throughout the supply chain. An assurance architecture that secures the destination of controlled technology while ignoring the provenance of its upstream inputs addresses only half of the integrity challenge.

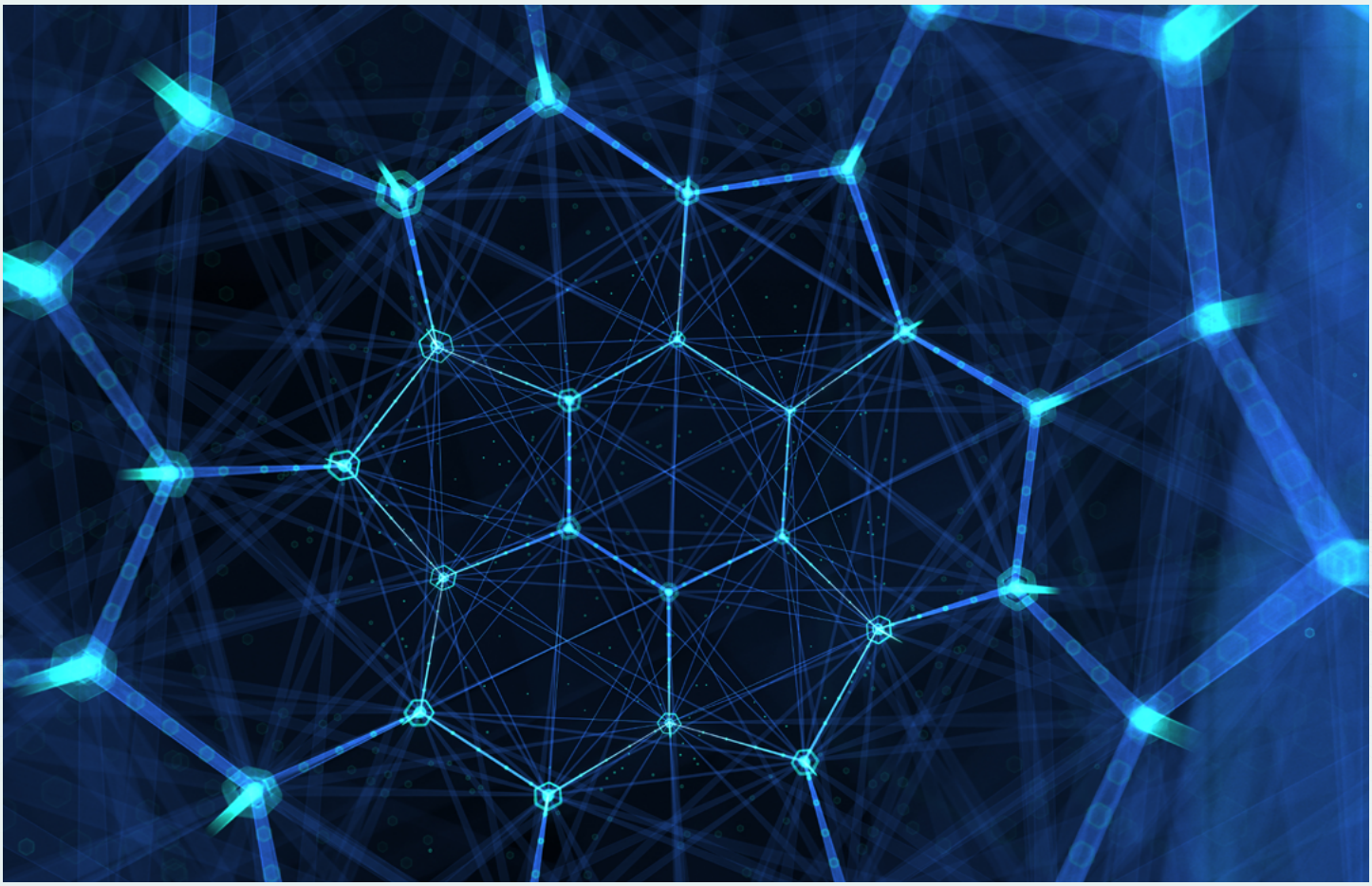
C. Symphonic Layers

The policy and implementation layers are symphonic. The coalition architecture creates the framework (the allied relationships, incentive structures, and governance norms) within which trusted deployments operate. The verification architecture makes those deployments credible and auditable. Neither works without the other.

An allied country that has signed on to a Pax Silica-compatible framework still needs an assured deployment architecture to demonstrate compliance with the conditions attached to US export licenses. A deployment that has impeccable verification architecture but no diplomatic framework supporting it operates in a governance vacuum that adversaries can exploit at the policy layer. The path to reconciling these imperatives runs through both layers simultaneously, and they reinforce each other: the policy layer expands the universe of trusted deployments, and the implementation layer makes each deployment more defensible and durable.

The verification architecture also creates a virtuous cycle of trust between trusted operators, allies, and the US government. Partners who engage proactively with the verification framework are investing not only in systems and infrastructure but in the trust conditions that enable faster action, less friction, and broader access in subsequent iterations. Each successful deployment cycle builds institutional confidence that reduces the cost and timeline of the next. This compounding trust dynamic converts compliance from a one-time qualification gate into a renewable credential, and it is one of the most powerful, and least discussed, strategic benefits of verification-by-design.

There are additional HOWs that the field is beginning to address: AI security research and publication standards, AI-enabled risk diligence tools for HPC environments, workforce development for export compliance professionals, and the emerging governance architecture for agentic AI systems that operate across jurisdictions. The June 3 convening was designed to surface and develop those additional dimensions. This paper establishes the foundation on which each of those branches can be built.



V. THE PATH FORWARD

A. What the Successor Regime Needs to Address

The following design principles represent the architecture toward which the policy and enforcement apparatus is converging, whether embodied in a formal successor rule, in the conditions attached to individual export licenses, in AAEP qualification standards, or in enforcement posture. A formal rule incorporating these verification-by-design principles would be the most efficient and durable expression of this architecture. These principles are actionable now through each of those channels, regardless of whether a rule materializes.

Outcomes-Based, Risk-Calibrated Qualification. Define what verification must demonstrate: access integrity, diversion prevention, use and output governance, risk-appropriate independent oversight, and allow operators to demonstrate compliance through architectures appropriate to their operating models. A research institution and a commercial cloud platform will implement those outcomes through different means; both can meet the standard if it is defined by what must be demonstrated rather than exactly how. Verification rigor should track risk signals rather than be applied as a uniform blanket. Large-scale compute deployments, sovereign programs, restricted-party-adjacent customers, and other red-flag indicators warrant the full architecture; routine commercial transactions in low-risk contexts do not require the same depth of scrutiny. Outcomes-based qualification operates at the qualification gate; the ongoing-verification layer, for higher-risk tiers, will necessarily inspect means as well as outcomes. The successor regime should specify a published evidence taxonomy that allows risk-appropriate calibration across the full spectrum of deployment contexts.

Policy Consistency and a "Level Playing Field" as a Design Requirement. Policy consistency is the single most important enabler of sustained investment in US AI infrastructure globally. Partners and investors need confidence that the rules governing access to American technology will be directionally predictable and durable across administrations and policy cycles. A regime that changes fundamentally every four years, or that is subject to retroactive or transactional reinterpretation, will increase friction, decrease reliance, and drive partners toward alternatives that offer greater predictability, even if technologically inferior. The verification-by-design architecture proposed here is intended in part to provide that durability: a program that has built and demonstrated assured compliance architecture has a credential that persists across policy cycles.

Full-Stack Coverage Across All Four Risk Dimensions. The four relevant risk dimensions are access, diversion, use and output, and verification. A rule that addresses only hardware location (the Chip Security Act's core mandate) is necessary but not sufficient for the governance challenges that operational deployments present.

End-Use and Access Visibility as a Structural Requirement. The successor regime needs to address the systematic visibility deficit that prevents the US government from understanding who is using controlled compute, for what purpose, and under what governance. This requires end-use reporting obligations that extend beyond the point of export to the point of actual use, including periodic attestation by operators and intermediaries, independent verification of end-use claims, and technology-enabled monitoring of access patterns. The intelligence and enforcement value of this visibility is significant: it would enable the United States to exercise the strategic leverage inherent in the global dependency on American AI compute in a targeted, informed, and intentional manner.

Enforcement Credibility. Adequate resourcing and commitment, extended statutes of limitation, penalty levels that reflect the severity of violations involving AI technology transfers, and consistent application will create an incentive and signal structure that promotes integration of verification-by-design principles into deployments and industry expectations and standards.

Compliance Liability Allocation as a Design Question. A verified large-scale AI deployment involves numerous parties: the OEM, the integrator, the operator, third-party oversight, the sovereign host, and BIS holding enforcement equities. Responsibility and liability for export control compliance need to be appropriately allocated among them, so that each party appropriately owns its exposure, and the capital community can price the resulting risk premium against the paper's claim that verified architecture reduces capital risk. The successor regime should specify how liability is distributed across the chain.

Intelligence Architecture as Foundational, Not Downstream. The verification capabilities the successor framework presupposes align closely with the institutional knowledge of the Intelligence Community rather than with rulemaking agencies. Treating inter-agency coordination as an implementation detail risks producing rules whose conditions exceed what current collection and analytic capacity can verify at scale. While technology-enabled verification can narrow the gap, it does not close it without the institutional architecture to act on what verification produces.

Interoperability Across US Licensing Instruments and Allied Frameworks. Controls lose force when allied jurisdictions adopt divergent definitions or parallel access pathways that restricted entities can route through. Qualification standards, KYC expectations, and assurance architectures should be developed with allied regimes from the outset rather than retrofitted into compatibility. Standards established for bilateral licenses should be consistent with consortium pathways such as the AAEP and compatible with the Pax Silica framework and emerging allied governance structures.

Adaptive Regime as Default. As successive generations of AI accelerators advance in efficiency and performance, and as frontier AI models continue to evolve, the export control regime that is fit for purpose in 2026 may quickly become outdated. Future verification technology and approaches will need to evolve to control who has access to advanced AI accelerators and limit what they can do with them depending on the recipient. The regime must be adaptive enough to confront the tradeoffs present in this highly dynamic area of innovation.

Trusted Verification as a Design Option. For the most complex and capable deployments, technology-enabled verification is necessary but not sufficient: the governance architecture that links human accountability, independent oversight, and government-validated reporting is what makes the system credible. Institutionalizing this role as a structural element of the AI deployment regime should include provider accreditation and performance criteria, conflict-of-interest controls, and market engagement to facilitate supply. This is not a call for additional regulation or for expanding the footprint of government agencies. The market is already producing the solution. Operators need to demonstrate compliance to access advanced compute; capital needs assurance to underwrite multi-decade infrastructure; the government needs confidence that its licensing conditions are being met. Independent verification and assurance, provided by qualified professional services firms, technology companies, and other market participants, is the market-appropriate response. These market participants can innovate faster and adapt to deployment-specific requirements in ways a government agency with a fixed mandate cannot. For the capital community, the analogy is direct: independent assurance functions like the independent assessment practices that govern complex transactions, infrastructure investments, and insurance underwriting, providing confidence, grounded in qualified professional judgment, that the deployment meets the required standard and the risks are appropriately managed.

B. Practical Considerations and Next Steps

Practical Implementation: Cost, Talent, and Time-To-Qualify. An honest framing of what assured implementation requires is essential for the capital, operator, and policy audiences this paper is meant to inform. A median large-scale AI program standing up a full assurance architecture from a cold start, sourcing the legal, security, audit, and regulatory liaison talent it does not have in-house, should expect approximately 12 to 36 months from program initiation to first verified production at scale. Total cost lands in the single-digit percent of program total cost of ownership, with the largest concentrations in initial architecture design, third-party engagement, and standing up the oversight and reporting capability described in Pillar Six. The binding constraint is talent: professionals fluent across export compliance, AI security, cloud architecture, and regulatory liaison are in scarce supply globally. The scaling of the model the paper proposes depends materially on how quickly that pool can be expanded.

Verification-By-Design as Operational Resilience. The operating environment described in this paper features binding supply chain constraints (memory shortages, component lead times, energy availability as the single largest physical bottleneck on new data center deployment) alongside escalating enforcement expectations. Operators who build assurance architecture from inception reduce their exposure to the exogenous shocks that disproportionately affect programs built without it: surprise enforcement actions, license revocations, counterparty failures arising from inadequate diligence, and regulatory changes that require costly retrofits. In an environment where supply chain disruptions are already compounding, the addition of a compliance failure on top of a supply chain event can be program-ending. Assurance architecture does not eliminate supply chain risk, but it materially reduces the probability and severity of the compliance-driven disruptions most within the operator's control. For the capital community underwriting multi-decade infrastructure, this operational resilience dimension is as significant as the compliance dimension itself.

Frontier Control Problems. The assurance architecture described in this paper was built for a workload mix dominated by training and inference for vetted customers and users. Two control problems will define the next three to five years and are not yet solved by the field. The first is the protection of trained model weights, against both static-file exfiltration and capability transfer through API distillation; hardware-attested location of compute is necessary but not sufficient, and the active control problem is who can query a model, at what rate, with what query characteristics, for what derivative purpose. The second is the governance of agentic AI systems, whose tool-calling and inter-agent chains break the observability assumptions on which conventional verification depends, and which current control standards (e.g., NIST AI 600-1, MITRE ATLAS, and OWASP Top 10 for LLM Applications) recognize as priority unsolved problems. These frontier items are priorities for the research and design agenda, and the assurance-by-design architecture must adapt controls appropriate to these emergent applications.

End-Use Verification as an Engineering Frontier. A central claim of the verification-by-design model is that end-use and access visibility can be meaningfully extended beyond point-of-export compliance into the life cycle of deployed AI systems. In practice, this remains only partially solved. Verifying end use at scale requires either inspection of inputs and outputs, which creates direct tension with tenant confidentiality and data-sovereignty regimes (including GDPR and Gulf data residency), or workload telemetry capable of characterizing use without accessing underlying content. Confidential computing and hardware-attested execution environments offer a promising pathway, providing cryptographic attestation that workloads are running in expected environments without exposing data. But removing the cloud provider from the trust equation with respect to data confidentiality does not remove the provider from responsibility for governance, access control, and use verification. The party that provisions compute, manages access, and controls the infrastructure retains obligations that hardware-level encryption alone cannot satisfy. Confidential computing strengthens the assurance architecture; it does not replace it.

"Vicarious use" risk, in which authorized customers provide services to downstream users, extends the verification boundary beyond the immediate operator in ways that are difficult to observe in real time. The implication is not that end-use verification is unattainable, but that it sits on an evolving engineering frontier. Within the most advanced verified deployment environments, practitioners are developing AI model control review processes designed to evaluate use, output, and access patterns at speeds human-centric review cannot sustain. In some architectures, this includes leveraging the high-performance compute infrastructure itself to develop and run AI-assisted compliance tools, turning the infrastructure that creates the verification challenge into part of the solution. These approaches are at varying stages of maturity, and the pathway to production-grade, at-scale use governance will require continued investment in both the underlying technology and the regulatory frameworks that authorize and constrain its use.

The successor regime will need to explicitly address this frontier, including how verification expectations interact with data-sovereignty law, commercial cloud architectures, and allied governance frameworks, and how AI-assisted verification tools are themselves governed, validated, and made auditable.

Forthcoming Assurance Architecture Annex. A companion implementation annex, developed in collaboration with practitioners from first-generation verified AI deployments and in dialogue with the standards community, will operationalize the verification-by-design framework into implementation-grade guidance. The annex will include: pillar-by-pillar control mappings to relevant US and AI-specific standards (NIST AI RMF 1.0 with the AI 600-1 Generative AI Profile, CSF 2.0, SP 800-53 r5, SP 800-207/207A, SP 800-63-4, SP 800-218/218A, and SP 800-234), along with software supply chain practices (SLSA, in-toto, SBOM in CycloneDX and SPDX); a threat model addressing diversion, remote-access exploitation, distilled-model exfiltration, and adversary-aligned use; alliance interoperability; independent assurance provider qualification, conflict-of-interest controls, and liability allocation; supply chain provenance; and the governance frontier for agentic AI and open-source dynamics. The annex is intended for presentation within the practitioner and standards community in fall 2026.

This paper's premise is that trusted partner relationships, combined with assurance architecture embedded in design and oversight from Day 1, are not policy friction or compliance overhead. They are what makes the entire framework credible, scalable, and defensible to US policy stakeholders, allied governments, technology providers, and the capital community simultaneously. They are the mechanism through which the United States can promote full-stack AI deployment at global scale and maintain assured confidence that the technology is not being exploited by adversaries. They are the architecture that will enable the United States to simultaneously run fast, deny adversary exploitation, and win the AI deployment race.

That is the resolution of the apparent dilemma, and it is already working: in the fastest supercomputer in the Middle East, in the world's leading Arabic large language model, and in the compliance architecture that BIS has referenced as the model for the next generation of export qualification standards. The work of the Verification-by-Design convening, and the deeper work that follows it, is to scale what is already working into the assurance architecture that governs the next decade of American AI leadership.

The companies and operators waiting for a specific rule to become law before building assurance architecture into their programs are waiting too long. The direction is clear. The instruments are converging. The time to build is now.

APPENDIX

About the Organizers

Alvarez & Marsal National Security, Trade and Technology (A&M NSTT) was the lead organizer of the Verifying the Stack convening.

A&M NSTT is a management consulting practice focused on the intersection of national security, foreign investment, trade controls, and technology security.

A&M NSTT serves as the US Department of Commerce-approved third-party security consultant for the KAUST Shaheen III GPU partition deployment and assisted in the design and implementation of the regulatory technology environment underpinning the G42 Assurance Compute Framework, which BIS has referenced as the model for future export license conditions. A&M NSTT has published extensively on AI export controls and submitted formal comments to the Department of Commerce on the AAEP RFI.

About the Authors

Randall Cook is a Managing Director and co-founder of the Alvarez & Marsal National Security, Trade, and Technology practice. A former Assistant United States Attorney and a Colonel in the US Army Reserve, he advises technology providers, defense industrial base contractors, sovereign AI operators, and financial sponsors on export controls, CFIUS, AI deployment governance, and the intersection of national security regulation with enterprise operations. He serves as the Department of Commerce-approved third-party security consultant on the KAUST Shaheen III GPU partition and has advised on the design of comparable frameworks in the Gulf and Indo-Pacific. A recognized voice on US AI export policy, the DOJ Data Security Program, and CFIUS enforcement, he has published in Westlaw Today, Reuters, WorldECR, and Corporate Compliance Insights.

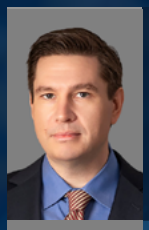
Kevin DeVilbiss is a Senior Director in the Alvarez & Marsal National Security, Trade and Technology practice, where he leads engagements at the intersection of AI deployment, export controls, and enterprise systems. His work spans export controls, compliance architecture design, and independent third-party assurance for large-scale AI deployments in the Gulf and beyond. Prior to A&M, Kevin was an executive at SAP Software, where he translated regulatory and business objectives into implementation-ready operating models, processes, and enterprise technology solutions. He was a moderator at the Verifying the Stack convening and anchors the practice's implementation-layer work with sovereign AI operators and technology providers.

Vince Mekles is a Managing Director and co-founder of the Alvarez & Marsal National Security, Trade and Technology practice. He specializes in complex and cross-border investigations, compliance advisory, and independent oversight, with deep experience translating regulatory exposure into operating model and control redesign for multinational enterprises. His work spans foreign investment review, sanctions and export controls, data security regulation, and the assurance architectures that make cross-border technology deployment durable. He is a co-author of A&M's published analysis of the DOJ Data Security Program and a recurring commentator on cross-border compliance. He anchors the practice's global compliance program and oversight work for foreign-owned suppliers, multi-jurisdictional manufacturers, and enterprises navigating overlapping regimes.

Louverture Jones is a Senior Director in the Alvarez & Marsal National Security, Trade and Technology practice, with nineteen years of cybersecurity, governance, risk, and compliance experience across the technology, aerospace and defense, energy, financial services, healthcare, and Department of Defense sectors. He is a deep practitioner across NIST 800-53, NIST 800-171, NIST AI RMF, FedRAMP, and CSA-CCM, with hands-on technology risk management experience in AWS, Azure, and hybrid cloud environments. Prior to A&M, he held senior positions at Black Rock Engineering & Technology, Deloitte, and Exelis. He anchors the practice's technical risk, cloud architecture, and CMMC controls work, and is a co-author of A&M NSTT's forthcoming implementation guidance for bifurcated enterprises.

AUTHORS

RANDALL COOK
MANAGING DIRECTOR, NSTT
randall.cook@alvarezandmarsal.com



KEVIN DEVILBISS
SENIOR DIRECTOR, NSTT
kdevilbiss@alvarezandmarsal.com

VINCE MEKLES
MANAGING DIRECTOR, NSTT
vmekles@alvarezandmarsal.com

LOUVERTURE JONES
SENIOR DIRECTOR, NSTT
louverture.jones@alvarezandmarsal.com

DISCLAIMER

This white paper is provided for informational purposes only and does not constitute legal advice. The views expressed are those of the authors and do not necessarily represent the views of Alvarez & Marsal or other convening sponsors or participants. References to specific enforcement actions, regulatory developments, and deployment programs are based on publicly available information as of the date of publication.

ABOUT ALVAREZ & MARSAL

Founded in 1983, Alvarez & Marsal is a leading global professional services firm. Renowned for its leadership, action and results, Alvarez & Marsal provides advisory, business performance improvement and turnaround management services, delivering practical solutions to address clients' unique challenges. With a world-wide network of experienced operators, world-class consultants, former regulators and industry authorities, Alvarez & Marsal helps corporates, boards, private equity firms, law firms and government agencies drive transformation, mitigate risk and unlock value at every stage of growth.

To learn more, visit: **AlvarezandMarsal.com**

Follow A&M on:

ALVAREZ & MARSAL
LEADERSHIP. ACTION. RESULTS.℠