

MAYER | BROWN

AI SUMMIT 2025

Thought Leadership Materials

Panel Session 5: Security Challenges



MAYER | BROWN

AI AND CYBERSECURITY

Legal and Policy Landscape

AI AND CYBERSECURITY

AI Threats

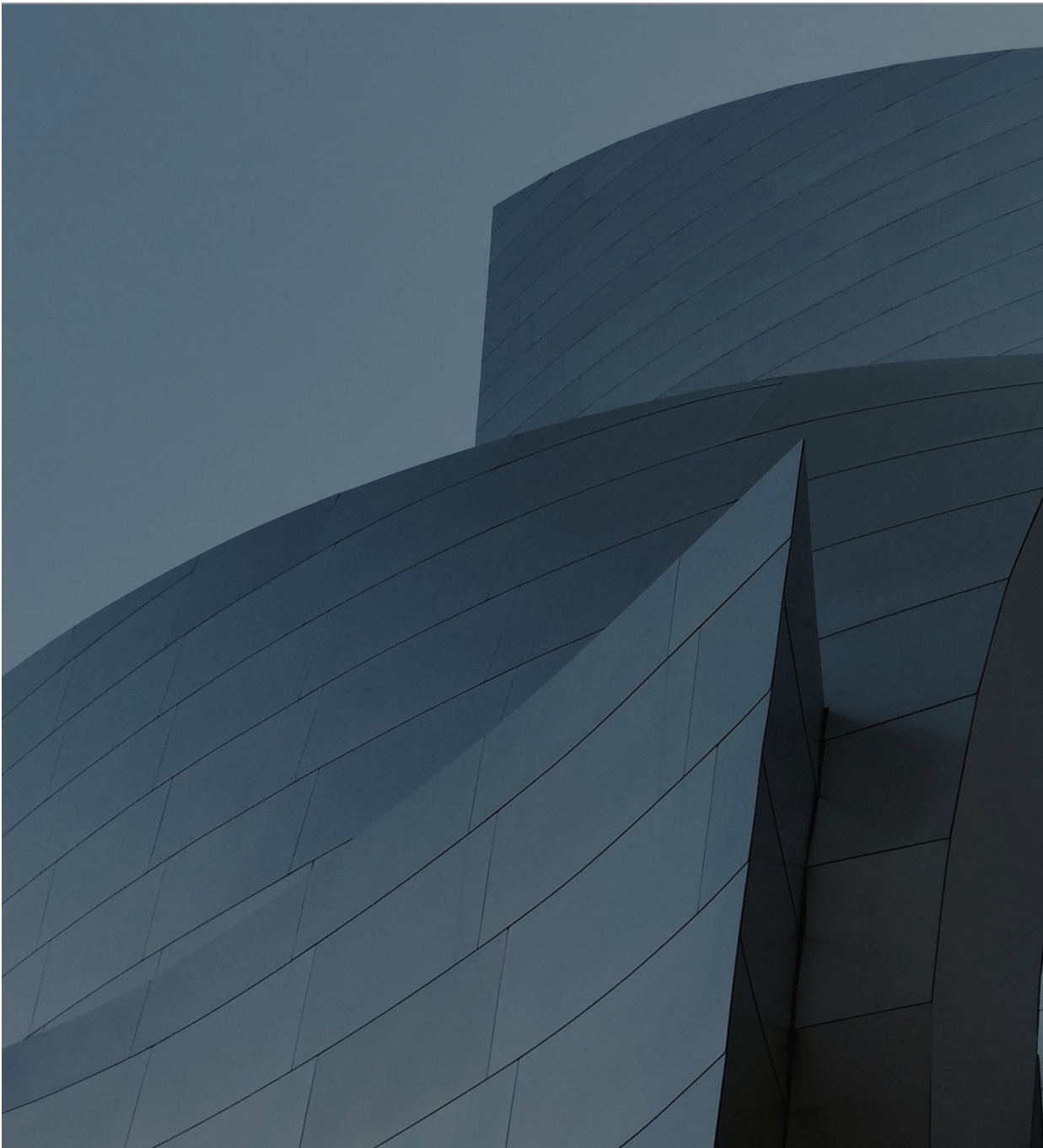
- AI-powered cyber attacks
- Attacks on AI

Securing AI

- AI Security
 - Expectations for developers
 - Expectations for deployers
- Red-teaming AI
- Responding to security incidents affecting AI

AI for Security

- Government support for use of AI for security
- Treatment of cybersecurity systems under AI regulations



NOT ON TODAY'S AGENDA:

- Non-cyber dimensions of AI safety (e.g., biological safety, chemical weapons, nuclear safety)
- Export controls
- Disinformation
- Algorithmic discrimination
- Online abuse
- Synthetic content

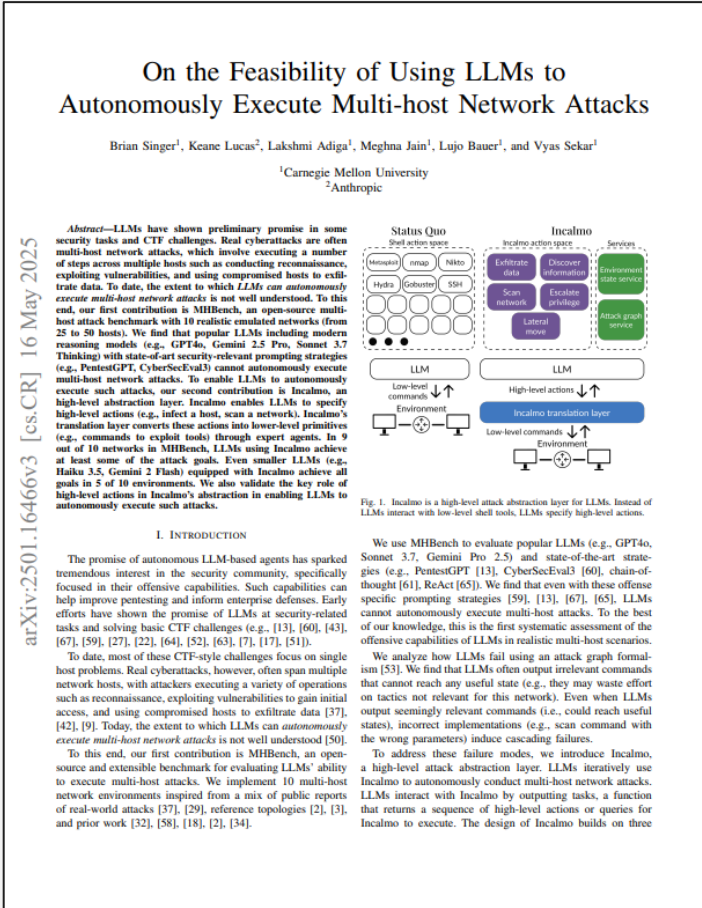


01

AI THREATS

AI-POWERED CYBER ATTACKS

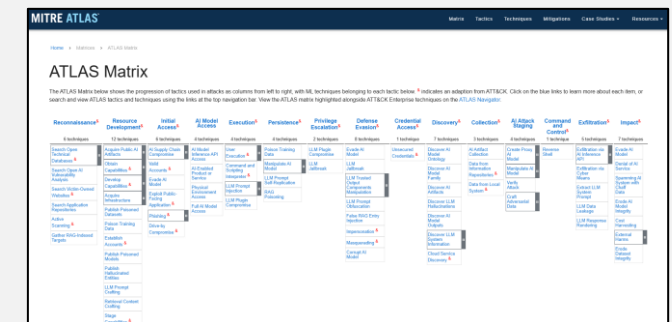
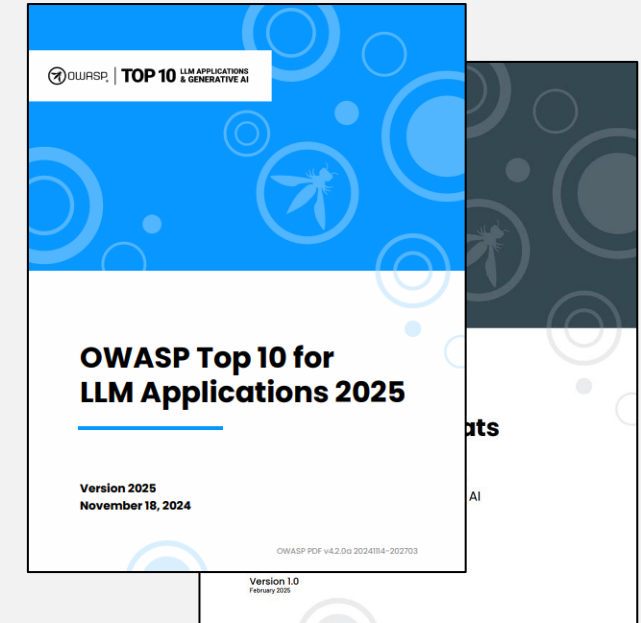
- Security teams and government officials have reported on the real-world use of AI to power cyber attacks, including through:
 - Deepfakes used in social engineering attacks;
 - AI-powered phishing campaigns;
 - AI-enhanced cybersecurity attacks (e.g., identify and exploit security vulnerabilities) and exploitation (e.g., perform reconnaissance, scan and analyze data).
- Abuse of agentic AI tools may further power these attacks.



Security researchers continue to [demonstrate](#) the potential for expanded malicious use of AI.

ATTACKS ON AI

- Policymakers are closely tracking the potential for a broad range of attacks on AI systems, including attacks that are common to other software-based systems and attacks that are distinctive to AI systems.
- Attacks include:
 - Evasion attacks: malicious input to fool the model or reduce its accuracy, e.g., prompt injection
 - Poisoning attacks, e.g., data poisoning, model poisoning
 - Information extraction attacks, e.g., model stealing, data reconstruction, membership or attribute inference attacks
 - Supply chain attacks, e.g., slopsquatting
- Companies can turn to an increasing number of resources to understand these attacks.





02

SECURING AI

AI SECURITY

- Policymakers have prioritized ensuring the security of the AI systems on which governments and businesses increasingly rely.
- Key focus areas for AI security include:
 - Data security
 - Application security
 - Model/model weight security
 - Infrastructure security
 - Securing AI output (code development)

The statistical, data-based nature of ML systems opens up new potential vectors for attacks against these systems' security, privacy, and safety, beyond the threats faced by traditional software systems.

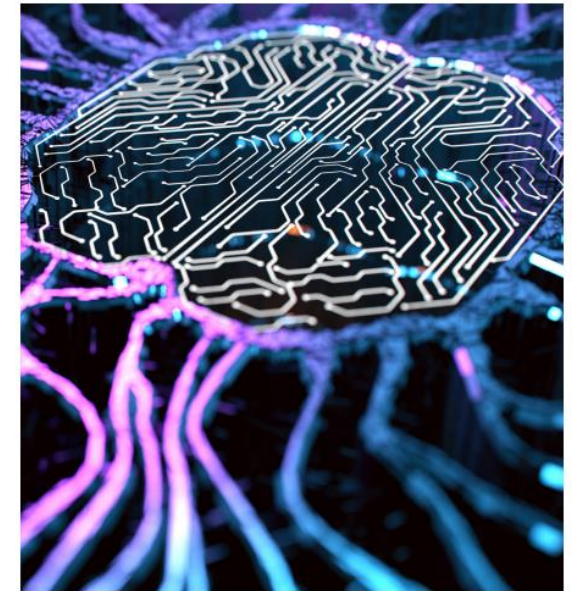
- NIST, Adversarial Machine Learning A Taxonomy and Terminology of Attacks and Mitigations (2025)



EXPECTATIONS FOR DEVELOPERS

- **General cyber risk measures:**
 - Secure SDLC, secure coding, and code review
 - Threat modeling, risk assessment, and vulnerability testing
 - Strong access controls and least privilege
 - Supply chain security and component provenance
 - Logging, monitoring, and incident response planning
- **AI-specific measures:**
 - Data provenance, integrity, and bias assessment for training data
 - Protection, versioning, and integrity of model weights and artifacts
 - Adversarial robustness testing, red teaming, and guardrails for prompt injection
 - Monitoring for model drift, data poisoning, and misuse
 - Documentation of model limitations, intended use, and failure modes
- **Considerations for the most powerful models**

Guidelines for secure AI system development



800
218A

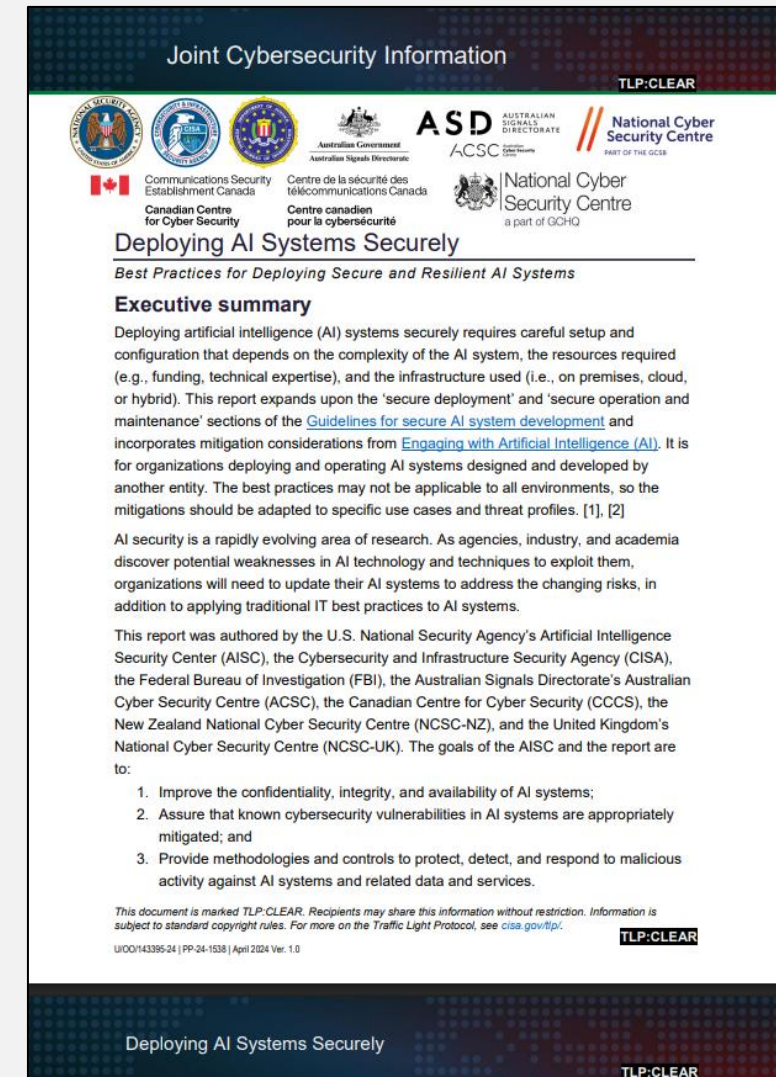
ces
Use
els
rofile

Booth
ppaya
assilev
Ogata
tanley
arfone

from:
0-218A

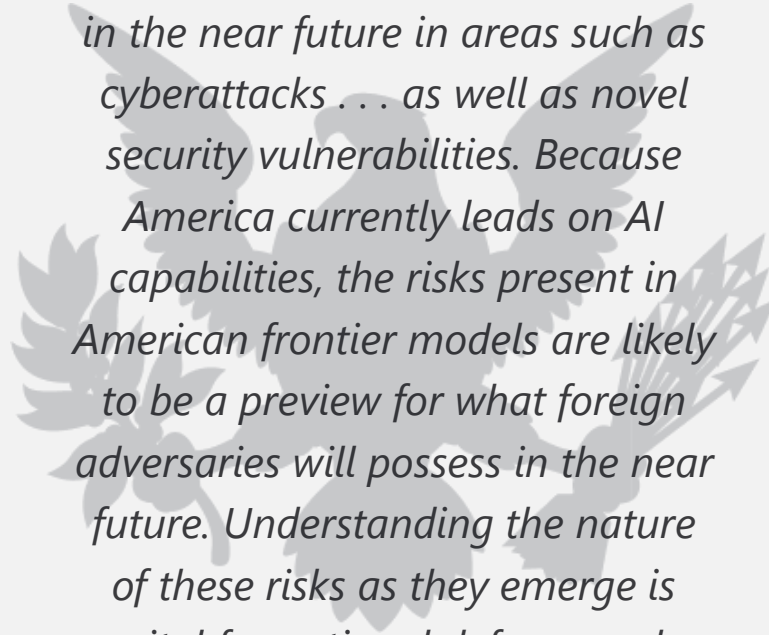
EXPECTATIONS FOR DEPLOYERS

- General cyber risk measures:
 - Establish robust governance and clear accountability
 - Conduct risk assessment and document threats
 - Harden configurations and keep systems patched
 - Secure APIs and use secure protocols
 - Promote security awareness, regular audits, and stay updated on emerging threats
- AI-specific measures:
 - Leverage threat models from AI system developers
 - Apply secure-by-design and Zero Trust to AI architecture
 - Encrypt and tightly control access to AI model weights and sensitive data
 - Validate AI artifacts' integrity and test models for vulnerabilities
 - Continuously monitor AI system behavior, inputs, and outputs



TESTING AI SECURITY

- **Distinctive aspects of AI red-teaming:**
 - Involves adversarial testing methods, e.g., attempts to elicit unwanted behaviors, subvert the model's built-in defenses or guardrails
 - Context-Dependent: Red-teaming practices and objectives vary by stakeholder (e.g., commercial developers vs. national security organizations) and by model type (general-purpose vs. specialized models)
- **Challenges:**
 - Measurement: what does it mean to "break" a model, and what constitutes a model failure or vulnerability?
 - Testing across multiple models and tracking results over time
 - Building consensus around testing practices and maintaining transparency
- **Particular questions for frontier models**



The most powerful AI systems may pose novel national security risks in the near future in areas such as cyberattacks . . . as well as novel security vulnerabilities. Because America currently leads on AI capabilities, the risks present in American frontier models are likely to be a preview for what foreign adversaries will possess in the near future. Understanding the nature of these risks as they emerge is vital for national defense and homeland security.

Winning the Race: America's AI Action Plan (July 2025).

RESPONDING TO AI SECURITY INCIDENTS

- Defining AI security incidents (vs. AI incidents)
- Distinctive features of AI security incidents:
 - Specific threat vectors, e.g., poisoned training dataset, supply chain attacks like malicious code that is executed when the model is loaded
 - Risk of compromise to sensitive and proprietary information, e.g., model weights, and to large datasets like training data
- Potential challenges ahead:
 - Identifying suitable remediation (e.g., in case of data poisoning)
 - Explainability of unintentional AI incidents, like algorithmic errors or system malfunctions
 - Complexity and impact of shutting off the model or AI system
 - Challenges relating to AI incident reporting and information sharing

EU Reporting Requirements

EU AI Act

For high-risk AI systems, mandatory reporting of serious incidents, but definitions are vague: *"an incident or malfunctioning of an AI system that directly or indirectly leads to the infringement of obligations under Union law intended to protect fundamental rights."*

Additional incident reporting obligations under **CRA, NIS2 and DORA.**

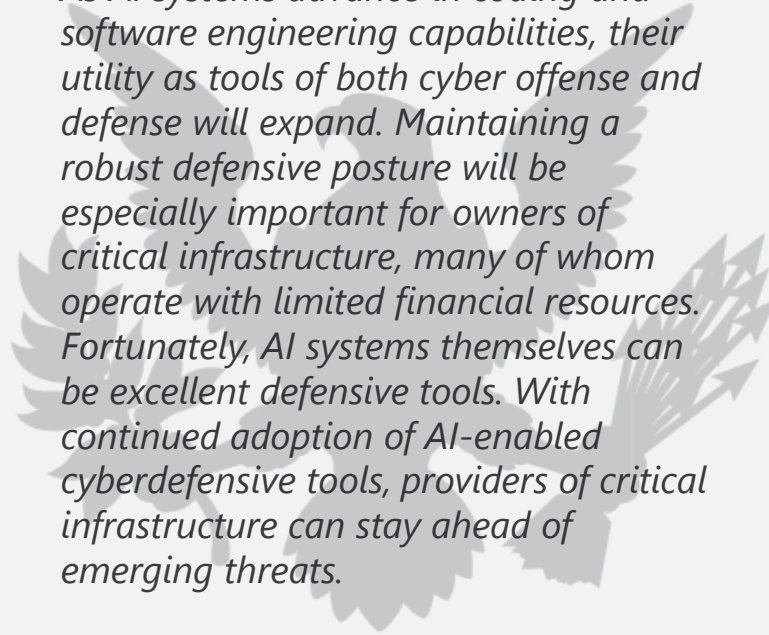


03

AI FOR SECURITY

AI FOR SECURITY

- AI promises to help companies make their defenses stronger and their incident response teams more effective, including through:
 - Vulnerability detection
 - Enhanced threat detection and response
 - Enhanced attack surface monitoring
 - Automated patching
- Governments globally have supported the use of AI for security to tip the balance toward cyber defenders
- Policymakers have evaluated how to avoid putting undue regulatory burdens on AI when used for security purposes



As AI systems advance in coding and software engineering capabilities, their utility as tools of both cyber offense and defense will expand. Maintaining a robust defensive posture will be especially important for owners of critical infrastructure, many of whom operate with limited financial resources. Fortunately, AI systems themselves can be excellent defensive tools. With continued adoption of AI-enabled cyberdefensive tools, providers of critical infrastructure can stay ahead of emerging threats.

Winning the Race: America's AI Action Plan (July 2025).



THANK YOU!

MAYER | BROWN

This Mayer Brown publication provides information and comments on legal issues and developments of interest to our clients and friends. The foregoing is not a comprehensive treatment of the subject matter covered and is not intended to provide legal advice. Readers should seek legal advice before taking any action with respect to the matters discussed herein.

Mayer Brown is a global legal services provider comprising associated legal practices that are separate entities, including Mayer Brown LLP (Illinois, USA), Mayer Brown International LLP (England & Wales), Mayer Brown Hong Kong LLP (a Hong Kong limited liability partnership) and Taill & Chequer Advogados (a Brazilian law partnership) (collectively, the “Mayer Brown Practices”). The Mayer Brown Practices are established in various jurisdictions and may be a legal person or a partnership. PK Wong & Nair LLC (“PKWN”) is the constituent Singapore law practice of our licensed joint law venture in Singapore, Mayer Brown PK Wong & Nair Pte. Ltd. Mayer Brown Hong Kong LLP operates in temporary association with Johnson Stokes & Master (“JSM”). More information about the individual Mayer Brown Practices, PKWN and the association between Mayer Brown Hong Kong LLP and JSM (including how information may be shared) can be found in the Legal Notices section of our website. “Mayer Brown” and the Mayer Brown logo are the trademarks of Mayer Brown. © 2025 Mayer Brown. All rights reserved.

MAY 16, 2025

US COMMERCE DEPARTMENT ANNOUNCES NEW EXPORT COMPLIANCE EXPECTATIONS RELATED TO ARTIFICIAL INTELLIGENCE

AUTHORS: THEA KENDLER, TAMER A. SOLIMAN, AIYSHA HUSSAIN, NICHOLAS T. JACKSON

On May 13, 2025, the US Department of Commerce's Bureau of Industry and Security ("BIS") unveiled heightened global due diligence requirements for companies using, granting access to, and trading in semiconductors used in artificial intelligence (AI). It also identified corresponding plans to remove worldwide license requirements on advanced semiconductors. While license requirements are expected to lessen under this announcement, BIS's expectations of AI industry compliance substantially increase.

BIS's guidance coincides with President Donald Trump's visit to the Middle East and significant new public commitments by US technology firms to build out AI infrastructure in the region. Although the details of a new regulation have not been released, together, these actions suggest the Trump Administration's willingness to encourage AI development outside the United States, while also expecting the AI industry to be significantly more attuned to end users and end uses.

BIS stated that it planned to rescind and would not enforce the worldwide controls on advanced semiconductors and AI model weights that President Joe Biden instituted in the waning days of his term. ([Read our Legal Update on the earlier rule](#)). License requirements would be maintained on select countries, including most Gulf states, but lifted for others, including India and Malaysia.

Together with this announcement, BIS released three guidance documents on expected due diligence associated with semiconductors, which outline:

- ***Due Diligence Guidance***: The risks of using semiconductors developed or fabricated in countries of concern, including China, anywhere in the world, including but not limited to Huawei's Ascend 910B, 910C, and 910D models, because of an inherent presumption that these semiconductors are subject to US jurisdiction;
- ***Diversion Guidance***: New "red flags" that may appear in a transaction, suggesting that illicit diversion of advanced semiconductors may be occurring; and
- ***Policy Statement on End-User and End-Use Restrictions for Training AI Models***: The potential enforcement consequences of providing access to advanced semiconductors and related items when the service provider knows, or has reason to know, that the items will be used to train AI models by or for parties headquartered in specific countries of concern, including China.

The key takeaway from BIS's guidance is that the US government expects the AI industry—including exporters, re-exporters, and data center operators—to conduct strict due diligence and screening to

prevent actions that are newly identified as violations of US law.

DUE DILIGENCE RELATED TO USE OF SEMICONDUCTORS

BIS's guidance advises that engaging in virtually any trade activity involving semiconductors developed or fabricated by companies located in, headquartered in, or whose ultimate parent company is headquartered in China or certain other countries of concern risks a violation of US export control regulations, and may result in substantial criminal and administrative penalties. Among other activities, this includes sale, transfer, export, re-export, financing, storage, and transport.

As a technical matter, the guidance broadly covers all semiconductors classified under Export Control Classification Number (ECCN) 3A090, and (in contrast to the guidance described below) is not limited to the "advanced" semiconductors in ECCN 3A090.a. Countries of concern include China, Macau, and all other countries in Country Group D:5 of the Export Administration Regulations ("EAR").

To reach this posture, BIS concludes that all such semiconductors "likely" fall within the jurisdiction of the EAR.

While BIS identifies Huawei's Ascend 910B, 910C, and 910D models as meeting ECCN 3A090's technical parameters and subject to its guidance, the agency does not limit its warning to these models.

To avoid exposure to a violation of the EAR, any party that seeks to take covered actions with respect to an ECCN 3A090 semiconductor may apply for a BIS authorization to engage in the proposed activity. Alternatively, if a party learns a violation has occurred that it was not involved in and does not otherwise have an interest in, it may submit a General Prohibition 10 waiver request.

BIS recommends confirming with reliable suppliers that a BIS authorization was in place covering the export, reexport, transfer (in-country), or export from abroad of both the semiconductor production technology from its designer to its fabricator, and the semiconductor itself from the fabricator to its designer or other supplier.

COUNTERING ILLICIT ADVANCED SEMICONDUCTOR TRANSACTIONS

In light of relaxed licensing requirements for advanced semiconductors, BIS also released updated guidance to increase the public's awareness of advanced semiconductor-related diversion schemes. Through a series of "red flags," BIS has identified new circumstances in a transaction that indicate the export, reexport, or transfer (in country) may be contrary to the regulations.

Significantly, if any such red flags appear in a transaction and are ignored, BIS may impose liability for a violation of the EAR. Ignoring a red flag may provide evidence of a "reason to know" that a violation of the EAR has occurred or is about to occur.

The newly announced red flags include, for example, if:

- the data center to which the advanced semiconductors or electronic assemblies does not or cannot affirm it has the infrastructure to operate the items;
- the delivery or installation address is unknown; and
- the customer is co-located with or its address is similar to a restricted party.

BIS further provided a list of due diligence steps that companies should take before conducting transactions involving advanced semiconductors and electronic assemblies with new customers,

especially those that are located outside of traditional US export control partner countries (i.e., Country Group A:1 of the EAR). These steps include:

- Before engaging in business with either domestic or foreign customers, notify such potential customers that your items are subject to the EAR and require a BIS license if exported, reexported, or transferred (in-country) to destinations for which a license continues to be required (i.e., Country Groups D:1, D:4, or D:5 (excluding destinations also specified in A:5 or A:6) of the EAR);
- Evaluate the customer's ownership structure to determine if parties are headquartered or have an ultimate parent headquartered in a destination in a country of concern, including China (i.e., Country Group D:5 and Macau); and
- Evaluate data centers to determine whether they have the infrastructure to operate electronic assemblies containing advanced semiconductors with power consumption greater than 10 megawatts. The guidance identifies that these data centers "merit additional scrutiny" because they may be capable of supporting high volumes of advanced semiconductors "for training AI models for or on behalf of parties headquartered in countries of concern, where such activities may support WMD or military-intelligence end uses/end users."

END USER AND END USE RESTRICTIONS FOR TRAINING AI MODELS

Through its policy statement, BIS has identified heightened expectations for the due diligence conducted by exporters, reexporters, and service providers into their customers and their customers' end uses. BIS announced that access to advanced semiconductors and other EAR-regulated commodities used for training AI models "has the potential to enable military-intelligence and weapons of mass destruction (WMD) end uses" in specific countries of concern, including China (i.e., Country Group D:5 and Macau). In line with this determination, BIS listed a number of activities that now potentially trigger a license requirement under the end-user- and end-use-based provisions of the EAR. 15 C.F.R. part 744.

The following activities may require a license when the provider knows or has reason to know that an AI model will be used for a WMD or military-intelligence end use/user:

- Provision of advanced semiconductors and commodities subject to the EAR when the exporter, re-exporter, or transferor knows or has reason to know that the recipient (e.g., a foreign Infrastructure as a Service ("IaaS") provider or data center provider) will use the items to train AI models for on behalf of parties headquartered in countries of concern, including China (i.e., Country Group D:5 and Macau);
- Changes of end use or end user of advanced semiconductors and commodities subject to the EAR, when there is "knowledge" that the transferee will use the items to train AI models for on behalf of parties headquartered in countries of concern, including China (D:5 countries or Macau); or
- A US person supports or performs any contract, service, or employment when there is "knowledge" such activity will be used for or may assist the training of AI models for or on behalf of parties headquartered in D:5 countries (including China) or Macau.

Persons conducting the activities listed above without a license are subject to potential civil or criminal enforcement action.

Finally, BIS notes that foreign parties acting contrary to US policy interests by training AI models that could support WMD or military-intelligence end use for, or on behalf of, parties headquartered in

Country Group D:5 may be listed on the Entity List.

RELATED ACTIVITIES

In addition to BIS's actions, congressional attention has increasingly focused on the AI industry in the last two weeks. Notably, two separate US Senate hearings examined, in part, the impacts of current US trade policy—including tariffs—on the domestic advanced semiconductor and AI sectors.

Further, companion bills introduced by a bipartisan group in the House or Representatives and Senator Tom Cotton, both entitled the Chip Security Act, would require location verification for advanced semiconductors, require that semiconductor manufacturers report and share information on potential diversion, and task the US Department of Commerce with analyzing additional steps to avert diversion.

Finally, a significant volume of AI industry trade was announced this week during President Trump's trip to Saudi Arabia, Qatar, and the United Arab Emirates. This week's BIS guidance indicates that AI data center development in the Middle East will continue to be subject to license requirements, although public announcement of these deals in concert with the President's visit suggests that such licensing will be expedited. To further this objective, certain countries ultimately may change Country Group designations in the EAR.

FINAL CONSIDERATIONS

BIS's pronouncements reflect a continued focus on preventing China from accessing AI technology, and the announced rescission of the worldwide advanced semiconductor license requirement may suggest that there will be an increased flow of trade in AI technology. However, the significantly increased due diligence requirements for the AI industry and service providers may ultimately lead to an onerous and uncertain process. One middle ground may be identified through the Validated End User program, which was instituted last year to facilitate global operations of trusted data center operators and service providers. Should the Trump Administration continue operation of this program, US hyperscalers and other trusted partners may determine that it provides a clearer route through many of these due diligence requirements.

Rapid changes in AI policy—along with the evolving US regulatory enforcement posture—present both risks and opportunities for the AI industry, and Mayer Brown is well positioned to advise companies in this dynamic sector.

AUTHORS

PARTNER

AIYSHA HUSSAIN

WASHINGTON DC +1 202 263 3051

AHUSSAIN@MAYERBROWN.COM

PARTNER

THEA KENDLER

WASHINGTON DC +1 202 263 3032

TKENDLER@MAYERBROWN.COM

ASSOCIATE

NICHOLAS T. JACKSON

WASHINGTON DC +1 202 263 3057

NJACKSON@MAYERBROWN.COM

PARTNER

TAMER A. SOLIMAN

WASHINGTON DC +1 202 263 3292

DUBAI +971 4 375 7160

TSOLIMAN@MAYERBROWN.COM

Mayer Brown is a global legal services provider comprising associated legal practices that are separate entities, including Mayer Brown LLP (Illinois, USA), Mayer Brown International LLP (England & Wales), Mayer Brown Hong Kong LLP (a Hong Kong limited liability partnership) and Tauil & Chequer Advogados (a Brazilian law partnership) (collectively, the “Mayer Brown Practices”). The Mayer Brown Practices are established in various jurisdictions and may be a legal person or a partnership. PK Wong & Nair LLC (“PKWN”) is the constituent Singapore law practice of our licensed joint law venture in Singapore, Mayer Brown PK Wong & Nair Pte. Ltd. More information about the individual Mayer Brown Practices and PKWN can be found in the [Legal Notices](#) section of our website.

“Mayer Brown” and the Mayer Brown logo are the trademarks of Mayer Brown.

Attorney Advertising. Prior results do not guarantee a similar outcome.

MAYER | BROWN

AMERICAS | ASIA | EMEA

MAYERBROWN.COM

Mayer Brown is a leading international law firm positioned to represent the world's major corporations, funds, and financial institutions in their most important and complex transactions and disputes. Please visit www.mayerbrown.com for comprehensive contact information for all our offices. This Mayer Brown publication provides information and comments on legal issues and developments of interest to our clients and friends. The foregoing is not a comprehensive treatment of the subject matter covered and is not intended to provide legal advice. Readers should seek legal advice before taking any action with respect to the matters discussed herein. Mayer Brown is a global services provider comprising associated legal practices that are separate entities, including Mayer Brown LLP (Illinois, USA), Mayer Brown International LLP (England & Wales), Mayer Brown (a Hong Kong partnership) and Tauil & Chequer Advogados (a Brazilian law partnership) and non-legal service providers, which provide consultancy services (collectively, the "Mayer Brown Practices"). The Mayer Brown Practices are established in various jurisdictions and may be a legal person or a partnership. PK Wong & Nair LLC ("PKWN") is the constituent Singapore law practice of our licensed joint law venture in Singapore, Mayer Brown PK Wong & Nair Pte. Ltd. Details of the individual Mayer Brown Practices and PKWN can be found in the Legal Notices section of our website. "Mayer Brown" and the Mayer Brown logo are the trademarks of Mayer Brown.