Office of the U.S. Chief Technology Officer
Eisenhower Executive Office Building
1650 Pennsylvania Avenue
Washington, DC 20504
Docket No. OSTP-TECH-2025-0067

Via Electronic Submission
October 24, 2025

Re: Request for Information; Regulatory Reform on Artificial Intelligence
COMMENTS OF THE MESSAGING, MALWARE AND MOBILE ANTI-ABUSE WORKING GROUP (M3AAWG)

I. Introduction

The Messaging, Malware and Mobile Anti-Abuse Working Group (M^{3}AAWG) appreciates the opportunity to comment on the Request for Information regarding the statutes, regulations, agency rules, guidance, forms, and administrative processes and other relevant restrictions that unnecessarily hinder the development, deployment, and adoption of artificial intelligence (AI) in the United States. We make these comments in our capacities as cybersecurity professionals and researchers committed to ensuring the security and stability of the internet.

M^{3}AAWG is a technology-neutral global industry association. With more than 200 members worldwide, we are the largest such organization in the online community. We bring together stakeholders in a confidential and trusted forum to develop best practices and cooperative approaches for combating online abuse. As a working body, we focus on operational issues of internet abuse, including technology, industry collaboration, and public policy. M^{3}AAWG works to fight online abuse caused by botnets, malware, spam, phishing, and ransomware.

II. Detailed Comments

Question i: What AI activities, innovations, or deployments are currently being inhibited, delayed, or otherwise constrained due to Federal statutes, regulations, or policies? Please describe the specific barrier and the AI capability or application that would be enabled if it was addressed. The barriers may directly hinder AI development or adoption, or indirectly hinder it through incompatible policy frameworks.

Answer i:

Thesis:

Current policy frameworks are hindering the detection of criminal behavior, allowing threat actors excessive access to U.S. network operations, and creating opportunities to victimize U.S. citizens. These threat actors, including foreign adversaries, are using advanced tools like AI to attack and compromise U.S. targets. Meanwhile, the U.S. is hamstrung in its ability to respond due to confusion around acceptable AI use, the lack of a national data privacy framework, and no clear safe harbor for information sharing.

The current patchwork of state-level privacy data regulations serves as a strong challenge to the smooth and confident handling, storage, processing, and disclosure of insights related to sensitive data. Security or anti-abuse models require diverse information sets (e.g., Distributed Denial of Service (DDoS), traffic patterns, routing data history, spam or phishing samples, URL and attachment feature sets, behavioral signals, complaint feedback, IP addresses, and domains) to effectively develop innovative AI-enabled countermeasures against abusive practices on the internet. This data often contains potentially sensitive information, yet it is instrumental to these efforts. Inconsistent privacy obligations and unclear allowances for secondary use (meaning operations model training and evaluation) can delay model improvements and discourage beneficial data stewardship, even when data is de-identified or aggregated. For example, if providers are hesitant to:

- Curate cross-provider data of abuse examples and payload signatures
- Train Large Language Model (LLM)-assisted classifiers on complaint and feedback loops
- Conduct red-team evaluations using synthetic or transformed samples derived from real attacks

The result is slower detection, more false positives, and decreased overall model efficacy.

Additionally, these unclear guardrails often lack safe harbors for sharing privacy-preserving abuse indicators, such as hashes of URLs, domains and senders; model-extracted features; and de-identified metadata, potentially creating legal risk and uneven collaboration. Government Information Sharing and Analysis Centers (ISAC) do not sufficiently replicate the communication mechanisms used by threat actors, nor do they provide rapid dissemination of Indicators of Compromise (IoCs) or adequate legal protections for participants. Organizations such as the National Institute of Science and Technology (NIST), with its AI Risk Management Framework, and the Organisation for Economic Co-operation and Development (OECD), through its AI Principles, have begun laying the groundwork to address the majority of these concerns. Adopting a federal AI governance framework could prove beneficial to the entire ecosystem, providing robust and easier-to-follow guidelines

that both protect Personally Identifiable Information (PII) and serve the greater public interest.

Question iii: Where existing policy frameworks are not appropriate for AI applications, what administrative tools (*e.g.,* waivers, exemptions, experimental authorities) are available, but underutilized? Please identify the administrative tools with specificity, citing the CFR or U.S.C. where applicable.

Answer iii:
We urge the Office of Science and Technology Policy (OSTP) to:

1. Coordinate cross-agency pilot programs that allow providers to evaluate AI-assisted abuse defenses under predefined safeguards, modeled on the FDA's "predetermined change control" approach to iterative updates.
2. Work with relevant federal government agencies to publish template conditions for privacy-preserving, pro-competitive sharing of abuse indicators.
3. Publish "no-action" criteria that agencies can adopt for limited, time-boxed trials of automated triage and takedown, paired with post-hoc sampling and drift monitoring.

These administrative tools already exist within the federal system (e.g., pilot and experimental authorities under 5 U.S.C. §301 and agency-specific innovation frameworks) and could be applied more consistently to AI. They also complement ongoing federal initiatives, including the Center for AI Standards and Innovation (CAISI) and NIST evaluations, and the Cybersecurity and Infrastructure Security Agency (CISA) and Joint Cyber Defense Collaborative (JCDC) playbooks, without weakening consumer protections. Instead, they accelerate measured, accountable, and innovative deployment of AI capabilities that serve the public interest and reinforce oversight, privacy, and fairness. They would complement ongoing federal initiatives, such as NIST and CAISI evaluations, and CISA and JCDC operational playbooks. Further, these approaches enable measured, accountable, and innovative deployment of AI capabilities that serve the public interest and reinforce oversight, privacy, and fairness.

Question iv: Where specific statutory or regulatory regimes are structurally incompatible with AI applications, what modifications would be necessary to enable lawful deployment while preserving regulatory objectives?

Answer iv:
We urge the OSTP to:

1. Modernize statutory and regulatory definitions that presuppose a human agent so that rules explicitly accommodate AI-generated communications *and* AI-driven defenses, while preserving consumer protection, for example, by defining the boundary between an AI agent and a human agent in AI-assisted activity.
2. Codify a security and anti-abuse carve-in for training and evaluation on content and metadata gathered in the ordinary course of operations, with strict retention, access, and audit controls.
3. Develop a model privacy-preserving data-sharing safe harbor that enables qualified entities to exchange AI-relevant abuse indicators (e.g., hashes, model embeddings, and reputation scores) while complying with existing privacy and competition law.

Question vi: Are there barriers that arise from organizational factors that impact how Federal statues, regulations, or policies are used or not used? How might Federal action appropriately address them?

Answer vi:
Fragmented mandates and inconsistent interpretation of AI-related policies across agencies create uncertainty and duplicative compliance efforts for providers. OSTP should establish an interagency coordination framework, modeled on JCDC or the Federal Privacy Council, to harmonize expectations for AI use in security and abuse-mitigation contexts. This group could build on existing efforts—such as the NIST AI Risk Management Framework (AI RMF), CAISI evaluations, and CISA playbooks—to promote consistent guidance and reduce regulatory friction.

Illustrative use cases and guidance needed:

1. General machine learning (ML) and AI-assisted tooling to aid in efforts, including but not limited to:
   - Spam detection
   - Phish, smish, and vish detection
   - Malware detection
   - ISP traffic analysis

   What we do: Train classifiers on labeled data, such as headers, payloads, user complaints, IP addresses, and traffic logs, with identifiers and features aggregated.
   The friction: It is unclear if using operational data for model training and evaluation qualifies as a permitted secondary use, whether the data is de-identified or identified.
   Guidance sought: Explicit recognition that personal data may be used for security and

anti-abuse training and evaluation, along with specificity on what levels of identification are allowed. Guidance should also include example controls, including hashing standards, retention limits, and auditability.

2. Cross-provider[1] and cross-vector abuse-signal sharing:
   What we do: Exchange hashed sending domains, URL fingerprints, signatures, and model features to identify abuse across many providers and vectors.
   The friction: There is legal uncertainty around "sharing/selling," joint controllership, and re-identification risk despite de-identification.
   Guidance sought: Model Memorandums of Understanding (MOUs) and a safe harbor for vetted entities that share privacy-preserving indicators for abuse prevention, with clear minimum technical and organizational safeguards.

General principles to consider in AI regulation

Ground any reforms in currently recognized frameworks such as the NIST AI RMF, OECD AI Principles, and the M^{3}AAWG AI Model Lifecycle Security Best Common Practices, emphasizing:

- Validity and readability
- Safety
- Security and resilience
- Accountability and transparency
- Explainability
- Privacy
- Fairness and bias mitigation

Encourage agencies to adopt or reference these reforms as non-regulatory tools for guidance and oversight activities. Greater clarity will improve interfirm and organizational collaboration, accelerating the development of novel anti-abuse solutions that better protect the public.

III. Conclusion

[1] Cross-providers can be identified as independent entities involved in the exchange or coordinated use of de-identified abuse-mitigation signals (e.g. mailbox providers, ESPs, ISPs, network operators, mobile carriers, messaging aggregators, registrars, and content delivery networks) and is not exhaustive of these categories.

Artificial intelligence offers immense opportunities but also significant risks. Like email and mobile before it, it can boost productivity, efficiency, and security when developed and deployed responsibly. As with earlier tools, however, it can also be abused and exploited by bad actors.

M[3]AAWG strongly supports the thoughtful advancement of AI systems that serve the public interest and protect individuals from harm. Clear, consistent, and privacy-conscious guardrails are essential, not to micromanage innovation, but to enable it to flourish safely. When rules are transparent and aligned with operational realities, organizations can act with confidence, collaborate effectively while offering a safe harbor for best-faith efforts, and innovate faster in the fight against online abuse.

We appreciate the opportunity to submit feedback and welcome further engagement as needed to answer any questions during this process. Please address any inquiries to M[3]AAWG Executive Director Amy Cadagin at comments@m3aawg.org.

Sincerely,
Amy Cadagin
Executive Director
Messaging, Malware and Mobile Anti-Abuse Working Group (M[3]AAWG)
comments@m3aawg.org
P.O. Box 9125, Brea, CA 92822