CYERA

June 10, 2025

The Honorable James Comer
Chair
Committee on Oversight and Government Reform
U.S. House of Representatives
Washington, DC. 20515

The Honorable Stephen Lynch
Ranking Member
Committee on Oversight and Government Reform
U.S. House of Representatives
Washington, D.C. 20515

The Honorable Nancy Mace
Chair
Subcommittee on Cybersecurity, Information Technology and Government Innovation
Committee on Oversight and Government Reform
U.S. House of Representatives
Washington, D.C. 20515

RE: Statement for the Record by Cyera, Inc. – Hearing on "The Federal Government in the Age of Artificial Intelligence" – June 5, 2025

Dear Chair Comer/Chair Mace/Ranking Member Lynch:

Cyera thanks you for the opportunity to submit comments for the record relating to the Federal Government in the Age of Artificial Intelligence. Founded in 2021, Cyera is the fastest growing data security company in the world and is a pioneer in the data security space. Cyera's customers use it to discover their data attack surface, control the use of data, monitor, detect, and quickly remediate risk. Deployable in minutes across any environment, Cyera's classifies sensitive data of all types without any human intervention thanks to its AI-powered classification engine.

Much like virtually every other entity across the globe, the U.S. Federal Government is working to rapidly adopt and utilize artificial intelligence (AI) across its information technology systems. AI brings with it the promise of revolutionary advances in efficiencies, cost savings, mission advancement, and taxpayer service. Successful AI systems can only fully succeed when they when utilize trustworthy data sets managed through robust data governance models.

In the case of AI use by the Federal Government, the solutions must have the ability to uniquely identify, classify, and control access to critical and sensitive data across complex, multi-cloud and on-premise environments. In an organization as large and complex as the Federal Government, that will obviously be an enormous challenge, but it can be met by utilizing advanced data intelligence platforms that offer automated, scalable solutions to implement data governance models. Done right, such governance models will enable agencies to secure their data, comply with regulatory mandates, and safely leverage AI for national priorities.

Data Intelligence Platforms

Modern data intelligence platforms employ automated discovery mechanisms that continuously scan and inventory data assets across all environments—cloud, on-premises, and hybrid. This includes structured databases, unstructured files, ephemeral storage, and software as a service (SaaS) applications. Automated mapping ensures agencies have a real-time, holistic view of where all sensitive data resides and thereby eliminates blind spots that can lead to unintentional exposure of sensitive data including personally identifiable information (PII), protected health information (PII), sensitive/secret data, and more.

The best of these platforms utilize AI-driven engines to analyze both the content and context of data:

- **Content Analysis:** Machine learning models inspect files and records for patterns indicating sensitive information such as PII, PHI, classified government records, and controlled unclassified information (CUI). The models adapt over time, improving accuracy as new data types and threat patterns emerge.
- **Contextual Analysis:** Beyond content, the platforms assess metadata, file locations, user roles, and business functions. For example, a document stored in a legal folder or accessed by a privileged user may be flagged as sensitive, even if the content is encrypted or obfuscated.

To balance speed and accuracy, the most effective solutions will combine multiple classification techniques:

- Metadata-Only Scans for rapid, broad visibility.
- Intelligent Sampling to quickly assess large datasets.

- Full Content Inspection for high-risk or highly regulated data.

This adaptive approach ensures precise classification without overburdening agency resources or impacting system performance.

Identity and Access Mapping

A critical component of data security is understanding who can access sensitive data. Advanced platforms integrate with identity and access management (IAM) systems to map user identities, roles, and permissions to specific data assets. This provides granular insight into which individuals, applications, or systems have access to each piece of sensitive information at any time.

Continuous Monitoring and Behavioral Analytics

Automated monitoring tools track all access events, helping to identify unauthorized behaviors. Behavioral analytics establish normal usage patterns and flag deviations, such as unusual data movement, off-hours access, or privilege escalation attempts. These alerts enable rapid investigation and response to potential breaches.

Integration with Security Operations

Data intelligence platforms integrate with security information and event management (SIEM) and security orchestration, automation, and response (SOAR) systems. This centralizes incident detection, investigation, and compliance reporting, streamlining the agency’s security operations and audit readiness.

Enabling Trustworthy AI

AI models are only as secure as the data they ingest. Ensuring that only authorized, non-sensitive, and compliant datasets are used in AI training and inference is essential to prevent data leaks, model bias, and regulatory violations. Automated classification and access control protect against inadvertent exposure of sensitive information through AI outputs, aligning with Department of Defense and National Security Agency guidance that AI system outputs must be classified at the same level as their inputs.

Supporting Regulatory Compliance

Federal agencies must comply with a range of mandates, including FISMA, FedRAMP, Executive Order 14110, and the AI in Government Act of 2020, which require robust data classification, access control, and auditability. Automated tools streamline compliance by providing continuous, verifiable evidence of data handling and access controls.

Mitigating Insider and External Threats

Continuous monitoring and access mapping are vital for detecting and responding to insider threats, compromised accounts, and external attacks. These capabilities are especially critical in the context of AI, where large-scale data aggregation and processing can amplify the impact of a single breach.

Operational Efficiency and Scalability

Automation reduces manual workload, eliminates human error, and allows security teams to focus on strategic risk management. Platforms designed for scalability ensure that agencies can maintain robust data governance as their data volumes and AI initiatives grow.

Alignment with Federal AI Governance Frameworks

Recent federal frameworks and guidance highlight the importance of data governance as a foundational element of responsible AI adoption. The Government Accountability Office's AI Framework and the National Security Memorandum on AI Governance emphasize the need for comprehensive data management, performance monitoring, and risk mitigation throughout the AI lifecycle. Automated data intelligence solutions directly support these principles by enabling:

- **Consistent Application of Data Policies:** Automated classification ensures that data policies are applied uniformly across all environments and data types.
- **Continuous Risk Assessment:** Real-time monitoring and analytics provide ongoing insight into data risks, supporting proactive mitigation.
- **Transparency and Accountability:** Detailed audit trails and access logs enable agencies to demonstrate compliance and accountability in AI operations.

Recommendations for Action

Even as Federal agencies work expeditiously to implement AI systems, it will be critical that specific governance steps are implemented as core requirements in order to enable their safe, secure, and efficient use. To that end, Cyera recommends that Federal agencies be required to undertake the following steps as part of the use of any AI system:

1. **Conduct a Comprehensive Data Inventory and Risk Assessment**
   - Action: Systematically inventory all data assets across the organization to understand what data exists, where it is stored, and its current usage and risk profile.
   - Rationale: Effective classification begins with visibility. Without a clear inventory, sensitive or regulated data may be overlooked, leading to compliance gaps and increased risk.
   - How: Engage stakeholders from privacy, security, compliance, etc., to define objectives and requirements. Use automated discovery tools to identify and map both structured, unstructured, and semi-structured data across all environments.

2. **Implement Automated, Persistent Data Classification and Access Controls**
   - Action: Deploy automated tools that continuously discover and classify sensitive data, and enforce access controls based on classification.
   - Rationale: Automation increases accuracy, scalability, and efficiency, reducing human error and ensuring ongoing compliance as data evolves.
     How: Integrate machine learning and AI-powered solutions to continuously scan, classify, and monitor data. Link classification to access management systems to ensure only authorized personnel can access sensitive data. Regularly review and update controls as new data is created or modified.

Automated, AI-driven data intelligence platforms are essential for the U.S. Federal Government’s secure and responsible adoption of AI. By uniquely identifying, classifying, and illuminating access to sensitive data, these solutions empower agencies to harness the transformative power of AI while safeguarding national interests, citizen privacy, and mission-critical operations. This capability is not just a technical advantage—it is a foundational requirement for the safe, secure, and compliant use of AI in the federal domain.

We thank you for the opportunity to submit the above comments and welcome further discussion as requested by the Committee and its staff.

Sincerely,

Lamont Orange
Chief Trust Officer
Cyera