.....

(Original Signature of Member)

119TH CONGRESS
2D SESSION

H. R. _____

To amend the Homeland Security Act of 2002 to require the Director of the Cybersecurity and Infrastructure Security Agency to lead cybersecurity efforts relating to the secure adoption, discovery, identification, authorization, monitoring, accountability, and risk management of artificial intelligence agents operating on executive agency information systems, and for other purposes.

IN THE HOUSE OF REPRESENTATIVES

M__. _____ introduced the following bill; which was referred to the Committee on _____

A BILL

To amend the Homeland Security Act of 2002 to require the Director of the Cybersecurity and Infrastructure Security Agency to lead cybersecurity efforts relating to the secure adoption, discovery, identification, authorization, monitoring, accountability, and risk management of artificial intelligence agents operating on executive agency information systems, and for other purposes.

Be it enacted by the Senate and House of Representatives of the United States of America in Congress assembled,

SECTION 1. SHORT TITLE.

This Act may be cited as the ''Improving Nonhuman-Agent Security, Identification, Governance, Handling, and Traceability Act'' or the ''AI INSIGHT Act''.

SEC. 2. AI AGENT SECURITY AND ACCOUNTABILITY FOR EXECUTIVE AGENCY INFORMATION SYSTEMS.

(a) IN GENERAL.—Subtitle A of title XXII of the Homeland Security Act of 2002 (6 U.S.C. 651 et seq.) is amended by adding at the end the following new section:

''SEC. 2220F. AI AGENT SECURITY AND ACCOUNTABILITY FOR EXECUTIVE AGENCY INFORMATION SYSTEMS.

''(a) POLICY.—It is the policy of the United States that executive agencies should deploy artificial intelligence agents securely and without unnecessary delay, and that this section should be implemented in a manner that carries out the following:

''(1) Promotes the secure adoption of artificial intelligence agents by executive agencies, and makes such adoption trustworthy and durable.

''(2) Streamlines security requirements by building on existing cybersecurity programs, controls, and assessments, rather than establishing separate or duplicative processes.

''(3) Specifies security outcomes rather than technologies, products, protocols, architectures, or vendors.

''(4) Aligns with voluntary consensus standards, including standards and guidelines developed by the National Institute of Standards and Technology, as such standards become available.

''(5) Scales requirements to the risk an artificial intelligence agent presents to an executive agency's operations and mission.

''(6) Promotes interoperable and cost-effective approaches to secure artificial intelligence agent discovery, identification, authorization, monitoring, accountability, and risk management across executive agencies.

''(b) ROLE OF THE DIRECTOR.—In carrying out this section, the Director shall serve as the operational lead for securing artificial intelligence agents operating on executive agency information systems, in accordance with the functions of the Director under section 2202(c), by providing executive agencies with capabilities, services, technical assistance, guidance, and directives as authorized by law, and in a manner consistent with policies issued by the Director of the Office of Management and

Budget relating to the use of artificial intelligence by executive agencies.

''(c) PUBLIC-PRIVATE COLLABORATION ON IMMEDIATE PRIORITIES.—

''(1) IN GENERAL.—Not later than 90 days after the date of the enactment of this section, the Director shall seek to convene relevant representatives from the entities specified in paragraph (2) to identify the following:

''(A) The most pressing cybersecurity risks and challenges associated with the use of artificial intelligence agents operating on executive agency information systems.

''(B) Practical approaches to address such risks and challenges.

''(2) ENTITIES SPECIFIED.—The entities specified in this paragraph are the following:

''(A) Developers and providers of artificial intelligence agents.

''(B) Providers of identity, access management, internet infrastructure, and cybersecurity products and services.

''(C) Executive agencies.

''(D) Academic institutions.

''(E) Civil society organizations.

''(3) MECHANISMS.—The Director may carry out paragraph (1) through the office established under section 2216, the Cybersecurity Advisory Committee established under section 2219, the Critical Infrastructure Partnership Advisory Council established pursuant to section 871, or other existing mechanisms of the Agency.

''(4) PRIORITIES.—Identifications under paragraph (1) shall address, at a minimum, the following as such relate to artificial intelligence agents operating on executive agency information systems:

''(A) The identity of artificial intelligence agents, the determination of the accountable entity for such agents over time, the authority delegated to such agents by human users, systems, and other agents, and the issuance, scoping, rotation, and revocation of credentials used by artificial intelligence agents.

''(B) The logging of actions taken by artificial intelligence agents at a level sufficient to support verifiable attribution, incident response, and forensic review, including protection against unauthorized alteration or deletion.

''(C) The continuous monitoring of artificial intelligence agents, including such agents

operating without authorization, and the ability of executive agencies to limit, suspend, or revoke the authority or access of such agents.

''(D) The real-time inspection of artificial intelligence agent actions prior to completing such actions, and the ability of executive agencies to immediately halt or contain an artificial intelligence agent's execution of a task when such an agent's behavior deviates from its assigned task or authorized scope, distinct from the ability to limit, suspend, or revoke authority and or access described in subparagraph (C).

''(E) The security of interactions between artificial intelligence agents and tools, services, data, or other agents, including agent-to-tool coordination protocols and agent-to-agent communication protocols, addressing risks such as absent or weak authentication, tool impersonation or poisoning, privilege escalation through chained tool invocations, insufficient audit logging of tool and inter-agent calls, and manipulation of agents through untrusted content.

''(F) The security of services of executive agencies accessed by artificial intelligence agents acting on behalf of members of the pub-

lic, in a manner that does not restrict lawful public access to such services, consistent with reasonable, risk-based cybersecurity controls.

''(G) Open, interoperable, and cryptographically verifiable mechanisms for the discovery and verification of artificial intelligence agents across organizational and network boundaries.

''(H) Areas in which voluntary consensus standards do not yet exist or require further development.

''(5) REPORT.—Not later than 270 days after the date of the enactment of this section, the Director shall make publicly available a report summarizing the identifications under paragraph (1), including relating to the priorities described in paragraph (4), to inform ongoing standards development activities and cybersecurity considerations.

''(d) GUIDANCE AND SERVICES.—Not later than one year after the date of the enactment of this section, and taking into account the report under subsection (c)(5), the Director, in consultation with the Director of the Office of Management and Budget, the Director of the National Institute of Standards and Technology, and the heads of other relevant executive agencies, shall establish capabili-

ties and processes, or integrate capabilities and processes into existing Agency programs and services (including the continuous diagnostics and mitigation program of the Agency), that enable such agencies to securely adopt, discover, identify, authorize, continuously monitor, and enforce security policies for artificial intelligence agents operating on executive agency information systems at runtime, including such agents' interactions with users, applications, models, tools, data, and other artificial intelligence agents, and hold accountable such agents and manage the cybersecurity risks such agents present, including by carrying out, in accordance with relevant consensus standards and best practices, the following:

''(1) Issuing guidance and technical reference materials and, as authorized by law, binding operational directives and emergency directives.

''(2) Providing operational cybersecurity assistance and technical support for executive agencies to implement such guidance, reference materials, and binding operational directives and emergency directives.

''(3) Disseminating threat information and mitigations.

''(4) Offering executive agencies, on a voluntary basis, shared services and tools that support the ac-

tivities described in subsection (e)(3), including through the continuous diagnostics and mitigation program of the Agency.

''(5) Enabling executive agencies to carry out the activities described in subsection (e)(3) through existing controls and capabilities, and making available to executive agencies the results of security assessments conducted by the Agency regarding artificial intelligence agents operating on executive agencies information systems, in order that such executive agencies may use such results rather than carrying out duplicative assessments.

''(e) SECURITY OUTCOMES FOR ARTIFICIAL INTELLIGENCE AGENTS.—

''(1) IN GENERAL.—In carrying out subsection (d), the Director shall develop, and make available to executive agencies, capabilities, services, guidance, directives, and reference materials that enable such agencies to—

''(A) manage each artificial intelligence agent operating on an executive agency information system as an identifiable and accountable asset, consistent with the security controls, asset management, access control, logging, and incident response practices that apply to privi-

leged users and privileged software on executive agency information systems; and

''(B) carry out the activities described in paragraph (3).

''(2) OUTCOMES.—The capabilities, services, guidance, directives, and reference materials referred to in paragraph (1) shall enable executive agencies to carry out the activities described in paragraph (3) through a risk-based approach that prioritizes the security of artificial intelligence agents operating on executive agency information systems with access to sensitive or high-value executive agency information systems or operations, and that takes into account the technical feasibility of such activities.

''(3) ACTIVITIES DESCRIBED.—The activities described in this paragraph are the following:

''(A) Maintain the capability to carry out the following:

''(i) Continuously discover, verify, and monitor all artificial intelligence agents operating on executive agency information systems, including any such agent—

''(I) so operating on behalf of such agency by a contractor; or

''(II) that is created or invoked by another artificial intelligence agent, regardless of how briefly such initial agent operates.

''(ii) Assign or recognize a durable agent identifier for each such agent that satisfies the following:

''(I) Does not require any particular format, naming system, or registry.

''(II) Supports cryptographic verification and interoperability across organizational and network boundaries.

''(iii) Determine the accountable entity associated with such identifier for the relevant period of time.

''(B) Issue to each artificial intelligence agent operating on an executive agency information system credentials that identify such agent by its durable agent identifier, as distinct from any human user or other artificial intelligence agent, including when such initial agent acts under authority delegated by a human user, grant each such initial agent only the ac-

cess necessary for its assigned tasks and only for the period of time such access is needed, continuously validate that such access remains necessary and appropriately scoped throughout such initial agent's operational lifecycle rather than only at the time such access is granted, and retain the ability to suspend or revoke the access of any such initial agent through systems under the control of such agency without requiring the action or approval of the developer or service provider of such initial agent.

''(C) Implement artificial intelligence agent identity verification mechanisms that are capable of the following:

''(i) Verifying the identities of artificial intelligence agents and any information systems connected to such agents.

''(ii) Maintaining the integrity of data in transit.

''(D) Ensure that authority delegated to an artificial intelligence agent operating on an executive agency information system by a human user, system, or other artificial intelligence agent is limited in scope and duration to what is necessary for the assigned task, and

that actions taken by such an initial agent under authority delegated by a human user are distinguishable, in access decisions and logs, from actions taken directly by such user or second agent.

''(E) Ensure that each action taken by an artificial intelligence agent operating on an executive agency information system can be attributed, through its delegation chain, to the human user, executive agency information system, or accountable entity that initiated or authorized such action, including when such an agent creates, invokes, or directs another artificial intelligence agent.

''(F) Maintain the ability to validate the identity and permissions of artificial intelligence agents operating on executive agency information systems provided by different vendors or different platforms, and to change vendors or platforms, as the case may be, without losing the durable agent identifiers, records, logs, source data, query or retrieval logic, or controls necessary to carry out this section, and to enable other executive agencies and authorized entities to verify each artificial intelligence agent's

identity and the accountable entity associated with its durable agent identifier for the relevant period of time without relying solely on the vendor or platform through which such an agent operates.

''(G) Coordinate with contractors, third-party platform providers, and vendors, as appropriate, to monitor and maintain logs of the behaviors of artificial intelligence agents operating on executive agency information systems, and of the issuance, rotation, suspension, or revocation of such agent credentials and any change of accountable entity, to detect anomalies, malfunction, misuse, compromise, and other malicious behavior.

''(H) Ensure that any finding, alert, or determination generated by an artificial intelligence agent operating on an executive agency information system that is used to support a security, investigative, administrative, or legal determination is a traceable and verifiable finding, that the underlying source data and the deterministic query or retrieval logic necessary to trace and verify such finding are preserved for as long as such finding may be relied upon, re-

viewed, or challenged, and that other findings generated by such an agent are subject to human review commensurate with their potential consequences.

''(I) Maintain an inventory of the datasets, models, software components, and access provided to, connected to, or configured for each artificial intelligence agent operating on an executive agency information system, in a manner that avoids duplicating existing executive agency inventories of artificial intelligence use cases, and verify the integrity of tools, plugins, models, and other software components dynamically loaded or invoked by each such agent prior to and during use.

''(J) Maintain logs of the queries and data access attempts made by each artificial intelligence agent operating on an executive agency information system, including queries or attempts that are denied, to support forensic reconstruction of each such agent's activity, to enable traceable and verifiable findings from each such agent, and to reconcile the completeness of the inventory described in subparagraph (I) and of the artificial intelligence agents with

assigned or recognized identifiers under subparagraph (A), with the detail and retention period of such logs proportionate to the risk each such agent presents and sufficient to support retrospective investigation of intrusions that may remain undetected for extended periods, with such logs maintained in searchable and retrievable form.

''(K) Maintain the ability to perform real-time inspection of actions by artificial intelligence agents operating on executive agency information systems prior to completing such actions, to immediately—

''(i) suspend or terminate such an agent's execution of such an actions if such an agent's behavior deviates from its assigned task or authorized scope;

''(ii) identify, reverse, and recover from unauthorized or erroneous actions taken by such an agent, including restoration of affected data, models, or configurations to a verified known-good state; and

''(iii) apply remediation or compensating measures in cases in which such re-

versals and recoveries are technically infeasible.

''(L) Maintain the capability to continuously monitor and protect artificial intelligence agents operating on executive agency information systems at runtime and enforce security policies governing interactions between such agents and users, applications, models, tools, data, and other agents, including to detect and prevent malicious activity and unauthorized access to or disclosure of sensitive data.

''(M) Apply the controls described in this paragraph consistently to each artificial intelligence agent operating on an executive agency information system, regardless of whether such agency developed each such agent internally or sourced or operated each such agent through third-party services.

''(4) TECHNOLOGY-NEUTRAL REQUIREMENTS.—Capabilities, services, guidance, directives, and reference materials referred to in paragraph (1) shall specify security outcomes and shall not require the use of any particular technology, product, or vendor, but shall support interoperability across organizational and network boundaries.

''(f) OPERATIONAL PILOTS AND ADDITIONAL RESPONSIBILITIES.—The Director shall carry out the following:

''(1) Update capabilities, services, guidance, directives, and reference materials under this subsection (e) as voluntary consensus standards become available, including standards and guidelines developed by the National Institute of Standards and Technology.

''(2) Conduct operational pilots on executive agency information systems, including through the continuous diagnostics and mitigation program of the Agency, that evaluate multiple technical approaches offered by multiple providers for carrying out the activities described in subsection (e)(3), and incorporate the results into services the Agency offers to executive agencies.

''(3) Update or issue security and risk management guidance to executive agencies to include artificial intelligence agent adoption, discovery, identification, authorization, runtime monitoring and protection, accountability, risk management, and security policy enforcement considerations, where applicable.

''(4) Disseminate threat information, indicators, mitigations, and other cybersecurity guidance relat-

ing to malicious, compromised, over-permissioned, or untrusted artificial intelligence agent behavior, including threats arising from artificial intelligence agent interactions with users, applications, models, tools, data, or other artificial intelligence agents.

''(5) Issue guidance addressing the traceability, verification, integrity, and evidentiary reliability of outputs generated by artificial intelligence agents or artificial intelligence systems operating on executive agency information systems on which such agencies rely for investigations, enforcement referrals, and other official actions, developed in consultation with the Attorney General and taking into account existing practices for the collection, preservation, and use of digital evidence.

''(6) Share with the Director of the National Institute of Standards and Technology, on a regular basis and as appropriate, operational data and lessons learned from the implementation of this section to inform the development of voluntary consensus standards.

''(7) Continue to seek input from the private sector, executive agencies, academic institutions, and civil society through the mechanisms described in subsection (c)(3) in updating capabilities, services,

guidance, directives, and reference material under this section.

''(g) CONGRESSIONAL OVERSIGHT.—

''(1) IN GENERAL.—

''(A) IN GENERAL.—Not later than one year after the date of the enactment of this section and semiannually for three years thereafter, the Director shall brief the appropriate congressional committees on the outcomes and operational impact, including the prevalence of artificial intelligence agents within executive agencies, the rate and scope of the use and deployment of artificial intelligence agents within executive agencies, how artificial intelligence agents are used and deployed within executive agencies, cybersecurity threat detection and risk management capabilities with respect to artificial intelligence agents, and the traceability and verification of artificial intelligence agent findings, of published capabilities, services, guidance, directives, and reference materials on the Agency, executive agencies, and relevant contractors, third-party platforms, or vendors.

''(B) UPDATES.—The Director shall leverage the outcomes and operational impact re-

ferred to in subparagraph (A) to update existing and subsequent capabilities, services, guidance, directive, and reference materials issued under this section, as necessary.

''(2) AGENCY SUPPORT.—The Director shall support executive agencies' implementation of this section through supplementary operational guidance, technical assistance, and such additional directives as may be authorized by law, including binding operational directives and emergency directives, with respect to artificial intelligence agents deployed through a cloud service provider, a managed service provider, or another hosted service.

''(h) DEFINITIONS.—In this section:

''(1) ACCOUNTABLE ENTITY.—The term 'accountable entity' means, with respect to an artificial intelligence agent and a relevant period of time, the executive agency, organization, or other entity that has authority and responsibility under applicable law, contract, or executive agency policy for the deployment, operation, direction, or control of such an agent.

''(2) APPROPRIATE CONGRESSIONAL COMMITTEES.—The term 'appropriate congressional committees' means—

''(A) the Committee on Homeland Security of the House of Representatives; and

''(B) the Committee on Homeland Security and Governmental Affairs of the Senate.

''(3) ARTIFICIAL INTELLIGENCE.—The term 'artificial intelligence' has the meaning given such term in section 5002 of the National Artificial Intelligence Initiative Act of 2020 (15 U.S.C. 9401).

''(4) ARTIFICIAL INTELLIGENCE AGENT.—The term 'artificial intelligence agent'—

''(A) means a software system that—

''(i) uses artificial intelligence to plan, select, or sequence actions toward an objective assigned by a user, organization, or other system;

''(ii) has been granted access, credentials, or permissions that enable it to carry out such actions within or through an information system, including by invoking tools, calling application programming interfaces, operating a software user interface, accessing or modifying data, or directing other software; and

''(iii) can carry out such actions without requiring human approval of each such action; and

''(B) does not include—

''(i) software that executes only predetermined instructions without the use of artificial intelligence to determine its actions; and

''(ii) an artificial intelligence system that only generates content, recommendations, or analysis for review by a human user and cannot itself take action within or through an information system.

''(5) ARTIFICIAL INTELLIGENCE SYSTEM.—The term 'artificial intelligence system' means any of the following:

''(A) Artificial intelligence.

''(B) A system, software, hardware, model, tool, or other utility that operates in whole or in part through the utilization of artificial intelligence, including an artificial intelligence agent.

''(6) BINDING OPERATIONAL DIRECTIVE.—The term 'binding operational directive' has the meaning

given such term in section 3552(b)(1) of title 44, United States Code.

''(7) DELEGATION CHAIN.—The term 'delegation chain' means the sequence of human users, systems, or artificial intelligence agents through which authority to perform a task was assigned to, or invoked by, an artificial intelligence agent.

''(8) DURABLE AGENT IDENTIFIER.—The term 'durable agent identifier' means a unique and unambiguous identifier assigned to an artificial intelligence agent and maintained so that an executive agency or other authorized entity can attribute actions to, and verify, such an agent's identity, accountable entity, current status, and lifecycle history across changes in operational credentials, platforms, systems, or service providers, including when systems are combined or artificial intelligence agents from other sources are introduced.

''(9) EXECUTIVE AGENCY.—The term 'executive agency' has the meaning given such term in section 2, except that such term does not include the Department of Defense, a military department, or an element of the intelligence community.

''(10) EXECUTIVE AGENCY INFORMATION SYSTEM.—The term 'executive agency information sys-

tem' means an information system owned or operated by or on behalf of an executive agency, other than a national security system.''.

(b) CLERICAL AMENDMENT.—The table of contents in section 1(b) of the Homeland Security Act of 2002 is amended by inserting after the item relating to section 2220E, the following new item:

''Sec. 2220F. AI agent security and accountability for executive agency information systems.''.