NATIONAL AI POLICY AND REGULATORY FRAMEWORK

Consumer Technology Association®

INTRODUCTION

This document sets out the elements of a regulatory framework intended to set guardrails for companies developing and deploying AI systems (the Policy). The Policy is intended to provide businesses flexibility to adopt AI risk management measures tailored to the specific risk profile of the AI systems they develop, deploy and/or implement.

Purpose. Given the rapid development of AI technologies, the Policy is intended to: (1) encourage appropriate guardrails and outcomes and (2) ensure that AI systems are safe, trustworthy, effective, ethical, and legal, rather than to focus on specific aspects or details of technologies.

Use of Existing Governance/Risk Frameworks. The Policy relies upon, and explicitly incorporates, elements of the NIST AI Risk Management Framework (NIST RMF), which was developed through a highly collaborative process with feedback from industry and other key stakeholders. Consistent with that, the Policy adopts a risk-based approach to regulation of AI systems, contains generally applicable AI governance requirements, and allocates certain responsibilities based on whether the business is a developer or deployer of the AI system.
Finally, the Policy recognizes that certain entities are already subject to sector-specific regulations and provides safe harbor protections for entities that have self-certified or obtained a third-party certification of compliance with an accepted AI risk management or governance standard.

SCOPING/DEFINITIONS/KEY TERMS

AI System. The Policy adopts the definition of "AI Systems" used in the NIST RMF: "an engineered or machine-based system that can, for a given set of objectives, generate outputs such as predictions, recommendations, or decisions influencing real or virtual environments" which are designed to operate with varying levels of autonomy. AI Systems can be purely software-based, acting in the virtual world (e.g., voice assistants, image analysis software, search engines, speech and face recognition systems), or AI can be embedded in hardware devices (e.g., advanced robots, autonomous cars, drones or IoT-based applications).

Developer. An entity is acting as a Developer of AI Systems when it codes, develops, or produces an AI System. It is possible that Developers will offer AI Systems to other entities for them to deploy.

Deployer. An entity is acting as a Deployer of AI Systems when it uses an AI System to interact with end users, or when it uses the outputs of an AI System to make decisions impacting end users. For example, where an organization licenses an AI system and puts that system (or capabilities) into the market, it would be a Deployer. [1]

Implementer. Entities or individual end users that implement or incorporate functionality and outputs of AI Systems for their own internal, or potentially external, uses.

High-risk AI Systems. AI Systems that are developed and deployed to effect certain outcomes that present heightened risks to individuals, communities or others. High-risk AI Systems are those which, at the time of development or deployment (either by licensed users or by third parties that modify the parameters of the systems) are based solely on: (1) automated processing and (2) produce decisions that have legal or equally significant effect on individuals, or which may impact individuals' health and safety.

[1] Although different obligations attach to each classification, it is possible for a single entity to be both a Developer and Deployer of AI Systems, depending on context. For example, where an organization codes, develops, or produces AI Systems and puts that system (or capabilities) into the market, it would be both a Developer and Deployer, and may be subject to several duties under each classification.

Application of the Policy to Small and Medium Sized Businesses. The purpose of the Policy is to create an ecosystem of uniformly trustworthy AI systems upon which consumers can rely. As a result, the Policy applies to all organizations (large and small) because the outcome and potential risk of noncompliance to the public is the same regardless of the size of the firm. Small and medium sized businesses may, however, require different treatment in the application of the policy to address differences in revenue, time in operation, and size of the user base.

Similarly, to avoid creating a potentially conflicting network of AI regulations, the Policy should not conflict with existing laws or regulations.

LEVERAGING EXISTING LAW

The Policy leverages existing law and standards that already govern AI System applications and outcomes. In certain instances, existing law already guards against potential bias and discrimination, regardless of whether such harm is human or machine-generated. The new Policy recognizes where such laws provide existing remedies and procedures and avoids duplication of the same.

TECHNOLOGY NEUTRALITY

The Policy focuses on applications and intended use cases, rather than the type of underlying model, algorithm or system that may be used. New guardrails should be principles-based and focused on outcomes rather than on the technical inputs of AI systems. As necessary, such guardrails and principles may also be informed by the type of system at issue, e.g., general purpose AI Systems; application-specific systems (e.g., image recognition, chatbot predictive analytics, content recommendation systems); predictive analytics; expert systems; and other categories.

RISK-BASED APPROACH

Given the wide range of AI Systems and applications/use cases of such systems, the Policy adopts a risk-based approach to AI governance, with oversight requirements being tailored to the nature and level of potential risk that an AI System may present.

Consistent with that approach, the Policy focuses AI governance obligations only on high-risk AI systems making decisions: (1) based solely on automated processing and (2) which have consequential legal or equally significant effect on individuals, or which may impact individuals' health and safety (hereinafter "automated decisions producing legal or similarly significant effects on individuals"). Decisions that impact an individual's ability to obtain financial services, education, housing, healthcare, and other essential services such as food and water should also constitute decisions that have critical legal or equally significant effect.

GENERAL RISK MANAGEMENT MEASURES

The Policy incentivizes entities to self-regulate by complying with an accepted third-party framework for AI governance, including the NIST RMF or applicable ISO, IEEE, CTA or related standards. Concerning risk management, and as an illustrative example, and consistent with the foundational elements of the NIST RMF, the Policy requires all entities in the AI ecosystem to adopt AI governance measures that adequately map, measure, manage, and govern risks of using AI Systems. Specifically, the Policy requires AI System Developers and Deployers to adopt AI governance programs with the following attributes:

1. Map, Measure, Manage and Govern AI Risks
    a. *Map*: Establish context of AI System; categorize AI System; identify capabilities, goals, and expected costs and benefits; map risks and benefits for all components of the AI System; characterize impact to intended users or individuals, and where necessary, groups, communities, organizations and society.

    b. *Measure*: Identify appropriate methods and metrics for measuring AI risks; AI Systems are evaluated for trustworthy characteristics; mechanisms for tracking identified AI risks over time are established; feedback mechanisms are implemented.

   c. *Manage*: AI risks identified following the map and measure steps (above) are prioritized and addressed; strategies to maximize AI benefits and minimize negative impacts are developed and implemented; risks arising from use of third-party AI systems are managed; risk mitigation strategies (including response, recovery and communication plans) are documented and managed.

   d. *Govern*: Policies, processes, procedures regarding AI risk management are implemented; accountability structures are in place and key personnel are empowered and responsible for oversight; workforce diversity is prioritized; organizational teams consider and communicate AI risks; processes are in place for robust engagement with relevant AI actors; policies and procedures are in place to address AI risks arising from third parties.

2. Voluntary Impact Assessments Where Necessary to Map and Measure AI Risks
   a. As appropriate, impact assessments may be conducted at an early stage of development, prior to launch, or after significant changes to an AI System. Impact assessments should indicate whether the AI System should be classified as high-risk, i.e., whether the AI System makes solely automated decisions producing legal or similarly significant effects on individuals. Such assessments shall be used for internal purposes and only disclosed upon consent of all parties involved in developing or deploying the AI system.

   b. Impact assessments should identify the risks of using the AI system, which should be quantifiable, concrete, and evidence-based where possible (and subject to the specific application or use case). AI Systems that may present more severe risks may require more detailed impact assessments.

   c. Where applicable and appropriate, the impact assessment should include a clearly defined and acceptable range of appropriate variation for performance of the AI system across different demographic groups, including marginalized groups.

3. Mitigation of Bias and Other Established Harms
   a. Where a high-risk AI System presents a risk of disparate treatment of individuals based on a protected classification, entities must put in place reasonable measures to mitigate such disparities. Such risks will be identified and classified under a defined term but will not prohibit intended biases that are inherent in most algorithms or models. Further, where a high-risk AI System presents a risk of other identified and likely foreseeable harms, such as privacy, cybersecurity and/or human security, entities must adopt reasonable safeguards to mitigate against such risks.

ALLOCATION OF RESPONSIBILITIES BASED ON ROLE IN AI ECOSYSTEM

The Policy contains a general requirement for all actors operating within the AI ecosystem to consider conducting voluntary AI impact assessments, but more detailed governance requirements should be allocated based on an entity's role as either a developer or deployer (or both) of AI Systems.

Further, the Policy also accounts for a third category of actors: "Implementers," or end users that implement or incorporate functionality and outputs of AI Systems for their own internal, or potentially external, uses. Such actors may engage in certain activities that could increase potential risks for other persons or entities. For example, in the case of generative AI applications, end users of large language model systems can use the outputs of the model for many potential purposes. Many will be innocuous, but some purposes, for example, using these tools to enable novel, realistic phishing schemes, can create risks for potential victims. When Implementers modify system parameters or use the outcome of the existing system for purposes outside of what it was designed for and/or implement a system that will result in a high-risk application, they will be required to adopt certain risk mitigation obligations, and must also adhere to the duties of developers and deployers outlined in this policy and regulatory framework.

The Policy sets out obligations that are specific to Developers, Deployers and Implementers of high-risk AI Systems.

Developers of high-risk AI Systems should:

1. Be subject to data governance/privacy requirements, including confirmation that a training data set includes data from diverse sources (where appropriate), that Developers of AI systems have obtained the consent, as necessary, to process identifiable data for the purposes of training an AI system, and that Developers adhere to applicable privacy laws.
2. Disclose to Deployers if the AI System was trained using personal information or other sensitive data, to the extent the Deployer is subject to applicable privacy laws. Disclose relevant metrics regarding the categories of data used. Confirm that if personal data was used to train the AI System, where necessary, consumer consents required under applicable privacy laws were obtained to process data for the purposes of training an AI system, and that Developers adhere to privacy laws applicable to the Developer or Deployer. These duties may be limited by applicable data minimization standards and principles.
3. Provide to Deployers and/or Implementers a set of instructions that lists the intended use cases and limitations of the AI System.
4. Inform Deployers if the AI System has been tested to the extent possible for accuracy, robustness, and unintended bias, and that reasonable mitigation measures have been taken where appropriate.

Deployers of high-risk AI Systems should be obligated to:

1. Conduct due diligence on Developers' AI Systems prior to purchase that would enable them to conduct a risk assessment/impact audit and to make required disclosures to consumers.
2. Perform regular audits of performance of AI System to detect emergent bias or other harms.
3. For AI systems that perform outside the acceptable variance based on demographic information, Deployers of AI Systems must adopt measures to mitigate the discrepancy in performance. The mitigation measures can be tailored to the specific uses of the AI system and can include additional levels of human review or re-training of the AI system.
4. Upon reasonable request from an individual subject to a decision by the AI System that is solely automated and produces legal or similarly significant effects on individuals, Deployers of AI systems must provide interpretations of the system's decisions and produce insights about the causes of its decisions.
5. When an AI system produces a decision by the AI System that is solely automated and which has legal or similarly significant effects on individuals, and is interacting directly with consumers, Deployers of AI systems should disclose that fact prominently, if not otherwise obvious.
6. Establish mechanisms for addressing end user feedback relating to use of the AI System.
7. Offer end users the ability to opt out of decisions solely made by AI Systems, when (1) the AI System's output is used to make decisions with "legal or similarly significant effects" and (2) the AI System does not incorporate meaningful human oversight.

Under certain circumstances that may be defined as between the Deployer and the Implementer, these duties may be assumed by the implementing party.

EXPLAINABILITY, TRANSPARENCY AND CONSUMER DISCLOSURES

Deployers of high-risk AI Systems must provide plain language explanations of how the AI system was designed and how it operates and produce insights about the basis of its decisions. In certain situations, Developers and Deployers of AI systems may be required to provide information about what data was collected to train the AI system.

When an AI system is interacting directly with consumers (such as a chatbot or other case where it could be mistaken for a human) and engaged in high-risk applications, then Deployers of AI systems should disclose that fact, as appropriate.

SELF-REGULATION AND COMPLIANCE WITH VOLUNTARY STANDARDS

The Policy also incentivizes entities to self-regulate by complying with an accepted third-party framework for AI governance. As such, the Policy provides that entities will be deemed in compliance if they have obtained and published a third-party certification or self-certification of compliance with accredited national and international standards, including the NIST RMF or applicable ISO, IEEE, CTA or related standards, as applicable.

SAFE HARBOR

Developers and Deployers that have published a self-certification of compliance with appropriate risk management and governance standards, or received certification of same from third parties, are entitled to a safe harbor. Specifically, such entities will have an affirmative defense to a cause of action alleging a violation of the AI regulations arising from this Policy.

REGULATORY SANDBOXES

Because commercial applications for this technology are still relatively nascent, and regulators across the globe continue to educate themselves on the power and promise of this technology, the Policy provides for the use of regulatory "sandboxes" to educate the AI stakeholder community and enable innovative ideas focused on novel approaches to mapping, measuring, managing and governing risks. Similar concepts have been proposed in other jurisdictions and would provide opportunities for regulators, commercial enterprises and civil society organizations to collaborate on the most effective means of addressing common goals.

EXEMPTIONS

As noted above, to avoid an overly complex and conflicting regulatory framework, the Policy creates exemptions for entities already subject to regulations, including those in healthcare, financial services, automotive and mobility/transportation and other highly regulated industries. In addition, the Policy includes exemptions for military, defense or national security purposes and for scientific research and development.

OVERSIGHT AND REPORTING

The Policy will grant authority to the Federal Trade Commission (FTC) to promulgate regulations and publish compliance guidance relating to the obligations described herein. The FTC will also adopt processes by which Developers or Deployers may file voluntary impact assessments with the FTC. The FTC would be restricted from using information obtained solely and exclusively through a covered entity's disclosure of information to the Commission for any purpose other than enforcing the Policy.

FEDERAL PREEMPTION

Finally, the Policy will be enacted by Congress as a matter of federal law and would expressly preempt all state and local regulations in this area.