Consumer Technology Association™

1919 S. Eads St.
Arlington, VA 22202
703-907-7600
CTA.tech

September 29, 2022

Consumer Technology Association
Comments on
NIST AI Risk Management Framework: Second Draft

The Consumer Technology Association® ("CTA")®[1] respectfully submits these comments in response to the National Institute of Standards and Technology ("NIST") request for comments related to the second draft of its Artificial Intelligence Risk Management Framework ("Framework" or "RMF") and the companion AI RMF Playbook ("Playbook").[2] As set forth in CTA's comments on the first draft of the Framework, CTA supports NIST's effort to create a flexible and voluntary risk management framework that will help identify and address risks in the design, development, use, and evaluation of AI products and services across a wide spectrum of types, applications, and maturity of AI systems throughout the AI lifecycle and "offer guidance for the development and use of trustworthy and responsible AI."

General Comments on the Framework's Second Draft

1. *Reaffirm Value of Risk-based Analysis of Opportunities and Threats Presented by AI*

"The goal of the AI RMF is to offer a resource for improving the ability of organizations to manage AI risks to maximize benefits and to minimize AI-related harms to individuals, groups, organizations, and society." *Second Draft, pg. 16.* CTA agrees that risk assessments are context specific, "likely to change and adapt over time," and that "risk tolerances can be influenced by policies and norms established by AI system owners, organizations, industries, communities, or policy makers." *Second Draft, pg. 9.* CTA also supports NISTs acknowledgment that the "RMF equips organizations to define reasonable risk tolerance, manage those risks, and document their risk management process." *Id.*

"Attempting to eliminate risk entirely can be counterproductive in practice – because incidents and failures cannot be eliminated – and may lead to unrealistic expectations and resource allocation that may exacerbate risk and make risk triage impractical." *Second Draft, pg. 10.* In light of the varying nature of risk across AI systems, CTA agrees with NIST that

[1] CTA® is the tech sector. Our members are the world's leading innovators—from startups to global brands—helping support millions of jobs. CTA owns and produces CES®—the largest, most influential tech event on the planet.

[2] Artificial Intelligence Risk Management Framework: Second Draft; released August 18, 2022. Available here: https://www.nist.gov/document/ai-risk-management-framework-2nd-draft. AI RMF Playbook; released August 18, 2022. Available here: https://pages.nist.gov/AIRMF/

organizations should adopt "a risk mitigation culture" and allocate resources such that they "align to the risk-level and impact of an AI system," *Second Draft, pg. 15*, recognizing that AI shortcomings and risks are an inevitable part of the AI development process. Although the RMF Playbook acknowledges that anticipated benefits of an AI system should be weighed against the anticipated risks and costs, CTA believes that explicitly referencing the risk-based calculus within the body of the Framework and a discussion of risk-based approaches or standards applied to AI systems would be helpful.

2. *Distinguish Risk Management Functions as Between Organizations Developing AI and Those Organizations Using AI*

CTA agrees that "third-party data or systems can accelerate research and development and facilitate technology transition. They may also complicate risk measurement because the metrics or methodologies used by the organization *developing* the AI system may not align (or may not be transparent or documented) with the metrics or methodologies used by the organization deploying or operating the system." *Second Draft, pg. 8*.

It is precisely for this reason that "risk measurement and management can further be complicated by how third-party data or systems are used or integrated into AI products or services." *Second Draft, pg. 8*, and it is "the shared responsibility of all AI actors [that] should be considered when seeking to hold actors accountable for the outcomes of AI systems." *Second Draft, pg. 15*.

CTA suggests that NIST clarify that all parties involved in the AI System should ensure the systems they develop and deploy as standalone or integrated components are trustworthy. The AI RMF uses case profiles should include situations involving entities that build and deploy their own models, but there will be many situations where an entity acquires and uses an AI system developed by a third-party developer. NIST should consider delegating specific responsibilities to both AI developers and acquirers. This delegation of responsibility, while important in all procurement contexts, may prove to be especially significant for developers of general AI systems and those that purchase such systems as components in larger AI systems. NIST should also clarify that developers and providers of data sets should take the measures available to them to ensure that the elements of trustworthy and unbiased AI are present in their products.

3. *Recognize Certain Elements of Trustworthy AI, Such as Bias and Fairness, Are Contextual and May Vary Depending Upon Circumstance*

"Fairness in AI includes concerns for equality and equity by addressing issues such as bias and discrimination, … [and] systems in which biases are mitigated are not necessarily fair." *Second Draft, pg. 14*. Because there is no universally accepted concept of fairness, and because bias cannot be eliminated in all circumstances, the Framework should enable organizations to make contextualized decisions to ensure that steps taken to measure, map, and govern risks are reflective of unique circumstances presented in specific situations where AI is deployed. Indeed, because decisions concerning bias may require tradeoffs between affected interests and intended goals of the system, developers and users of trustworthy AI systems must be empowered to take

a contextual approach to risk assessment and management recognizing that acceptable risk will always be use-case specific.

4. Purveyors of AI Should Communicate Capabilities and Limitations of AI

The Framework does not currently include any recommendations that developers of AI systems which distribute their systems, either as finalized products or components of larger AI systems, design them to allow for further fine tuning using the acquirer's data. Similarly, while the Playbook directs acquirers to establish policies related to the limitations of third-party AI systems, it does not contain any direction for purveyors to proactively provide information regarding the limitations of their AI systems (and need to fine tune them) to acquirers. AI system developers should provide documentation regarding the limitations of the AI systems and the process to allow for those limitations to be mitigated.

5. Include Decommissioning in Lifecycle of AI Systems

In its comments on the initial draft, CTA recommended that NIST contemplate the decommissioning and phasing out of AI systems and offer baseline risk management considerations when phasing-out the use of AI systems. The Second Draft (and Playbook) did not address this.

CTA reiterates its recommendation that NIST address decommissioning and phasing out of AI systems in its efforts to offer a framework to mitigate risk across the entire lifecycle of AI systems and thus increase the likelihood that decommissioning of the AI system does not jeopardize the organization's trustworthiness – and that decommissioning does not increase risks and decrease trustworthiness.

6. Recognize That Transparency Tools Are Varied and Being Developed

The Framework should recognize that various AI transparency tools (e.g., system cards, model cards, etc.) are being developed, and it has not been established which method represents the best approach, or if certain methods are best suited to specific situations. Accordingly, developers of AI systems should be encouraged to test out different types of transparency tools and follow industry standards at the time a model is in use

Comments on the "AI RMF Core": Mapping, Measuring, Managing and Governing

1. Section 6.1 – Map Function

CTA reiterates its recommendation that the Map function address the selection and collection of data, and that NIST expand the data selection and collection components in Map 2 to specifically include data risk mitigation strategies such as: (1) mapping or inventorying data; (2) classifying data and sourced datasets; (3) determining possible sources of corrupt or misplaced data and data sets; and (4) analyzing risks associated with the data sets.

CTA encourages NIST to identify when and where—in the Map function or elsewhere—unintended *potential* consequences are actually identified for mitigation. While the Framework

includes consideration of positive and negative consequences, it is not clear if those include a process for anticipating unintended consequences. Indeed, Map 1.6 discusses integrating feedback about "unanticipated negative impacts," *Second Draft, pg. 22*, but provides no specifics about the "conditions and circumstances" that could lead to potential negative consequences. *Id., pg. 13*.

2. Section 6.2 - Measure Function

CTA appreciates NIST's recognition that risk should be measured and evaluated throughout the lifecycle of the AI system and encourages NIST to identify lifecycle stages at which risks could be reevaluated (*i.e.*, when a Framework user should reengage with the AI RMF Core). This is consistent with CTA's recommendation, above, that the Framework address AI system evolution over time.

Relatedly, CTA reminds NIST that risks and lifecycles are not the same for every algorithmic model created by a particular developer or generated for a particular purpose, since each model is built differently from others and based on different datasets. There are instances where risk cannot be measured. CTA respectfully asks NIST to provide guidance for those instances, including that the absence of an ability to measure risk does not imply that an AI system poses high or infinite risk. Additional clarity on these situations will ensure that the absence of measurement does not automatically or necessarily result in halting the development or use of a technology—or the implementation of misplaced or unnecessary mitigation measures under an incorrect assumption of high risk at a stage of the lifecycle where such measures would not be useful.

3. *Sections 6.3 & 6.4 - Manage and Govern Functions*

CTA reiterates its suggestion that the Framework better account for the differing responsibilities of AI system developers and end users. CTA suggests that NIST collaborate with industry stakeholders to develop additional guidance for allocating risks, responsibilities, and obligations between these two groups. What obligations belong to developers? Which belong to users? How should the two groups interact? How should AI system components be evaluated as part of a larger AI system? Each actor in the system development, operation, and modification cycle has a distinct insight into the potential risks and has attendant risk spotting and mitigation responsibilities.

As part of Govern 1.1, the Playbook provides that when auditing an AI system, organizations can document whether existing legislation or regulatory guidance been "reviewed understood, and managed." It would be helpful to get clarification on whether auditing is intended to be based on legislation and regulatory guidance, internal periodic review procedures, or some other principles.

As part of Govern 1.2, the Playbook provides that there be a process to determine whether "characteristics of trustworthy AI are integrated into organizational policies, processes, and procedures." Our concern here is that certain of this information would not be appropriate to make available as it could get into the hands of nefarious actors, especially those who may skew

external stakeholder feedback to unfairly change the AI system to their benefit, and that may not be shared with other AI actors and developers.

Conclusion and Additional Recommendation

CTA is encouraged by NIST's attention to the comments and concerns of AI users and developers and finds that the Second Draft exhibits measurable improvements, with some issues still needing consideration as set forth above.

Although NIST recognizes concerns related to privacy, safety, and infrastructure generally, CTA encourages NIST to consider explicit cybersecurity guidance in the Framework. Securing input data, models, and algorithms from tampering or unsupervised changes are necessary to further reliability (*e.g.*, that models produce anticipated outcomes) and protect against bad actors. Security governance should include ongoing audits and monitoring to confirm that systems behave as intended, have not experienced unauthorized internal access or modification, and enjoy robust security to avoid adversarial attack. Cybersecurity guidance should also address privacy, security, and infrastructure considerations related to sharing data and models, such as between stakeholders, between private-and-public actors, and with developers. Potential breaches or algorithmic corruption and unauthorized disclosures of personal information are not only an internal concern of AI users and developers, but the entire technology ecosystem.

Respectfully submitted,

/s/ Douglas K. Johnson
Douglas K. Johnson
Vice President, Emerging Technology Policy

/s/ Michael Petricone
Michael Petricone
Sr. Vice President, Government and Regulatory Affairs