Consumer Technology Association™

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

September 15, 2021

Consumer Technology Association
Comments on
RFI - NIST AI Risk Management Framework

The Consumer Technology Association® ("CTA")®[1] respectfully submits these comments in response to the National Institute of Standards and Technology ("NIST") request for information ("RFI") related to the NIST Artificial Intelligence Risk Management Framework ("Framework").[2] CTA applauds NIST's thoughtful work on these issues and the effort to establish a voluntary framework for incorporating trustworthiness considerations into the creation and use of artificial intelligence ("AI") products, services, and systems.

CTA particularly agrees with the proposed Framework attributes of adaptability, risk-based, outcome-focused, voluntariness, and non-prescriptiveness, that collectively will foster public confidence and trust. CTA agrees with NIST's statement that because there is no objective standard for ethical values, in part because such values are grounded in the norms and legal expectations of specific societies or cultures, there is a general consensus that "AI must be designed, developed, used, and evaluated in a trustworthy and responsible manner to foster public confidence and trust."[3]

Principles and frameworks, such as those being developed by NIST, will have lasting impact on the development of AI systems to the extent that stakeholders "buy-in." Uniform adoption of and respect for global AI-related principles and frameworks are most likely if these principles and frameworks encourage the development of accurate, ethical, inclusive, and trustworthy AI. At the same time, these policies must offer voluntary, flexible oversight and compliance solutions, while retaining the essential economic incentives to innovate.

AI technology offers tremendous opportunities for human and societal development, and CTA agrees that ***AI must be trustworthy***. AI can promote inclusive growth, improve the welfare and well-being of people, and enhance global innovation and productivity. It has profound promise for our national interests, as well. Indeed, as Executive Order 13859 states, "[c]ontinued American leadership in AI is of paramount importance to maintaining the economic and national security of the United States and to shaping the global evolution of AI in a manner consistent

---

[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, Docket 21076-01510, 86 Fed. Reg. 40810, National Institute of Science and Technology, Department of Commerce (rel. July 29, 2021) (hereafter "Framework" or "AI RMF").

[3] *Id*., 86 Fed. Reg. at 40810.

with our Nation's values, policies, and priorities."[4] Doing so requires trust in systems: from policymakers, regulators, users, and the public. The Framework can help build this trust.

Because trust is inescapably linked to perception of risk, the Framework must enhance efficiency, effectiveness, and user satisfaction that are tied to perceived technical trustworthiness. Of course, technical trustworthiness varies with perceived and actual risk factors: for example, we know that industrial AI systems working in tandem with humans in factories, as well as medical and diagnostic AI systems, may present higher risks for personal safety and well-being, while music and media recommendation AI systems likely present very low risk. The Framework must recognize these distinctions based on both perceived and actual risks.

Further, the utility of frameworks like that which NIST is developing is that they can serve to complement developing policy initiatives and encourage the use of self-regulatory practices, through the adoption of codes of conduct, voluntary standards, and best practices. When based upon clear and targeted frameworks, self-regulation can result in meaningful protection for users and profound innovation while enhancing trustworthiness. NIST has an opportunity with this Framework to champion such an approach. To that end, CTA is encouraged that NIST recognizes the Framework must be "risk-based, outcome-focused, voluntary and non-prescriptive."[5]

Respectfully, CTA urges NIST to consider opportunities to strengthen the Framework by reviewing its application in practice. This is necessary to better understand and evaluate the potential costs, and benefits, of implementing a framework that may lead to new rules and restrictions on the development, use, and sale of AI.

In the RFI, NIST requested feedback on twelve specific topics. CTA offers its perspective on several of these.

I. How organizations currently define and manage principles of AI trustworthiness and whether there are important principles which should be considered in the Framework besides: transparency, fairness, and accountability (RFI No. 3)

CTA suggests NIST consider these additional principles for the Framework, each of which supports the development of trustworthy AI:

Ability to implement. The Framework should emphasize that its principles and standards must be more than theoretical concepts but must also be implementable. The success of the Framework will hinge on whether it is achievable and can be implemented in a meaningful way. To accomplish this, CTA highlights that the Framework should be attentive to *fact-based risk assessment and remediation processes*. A fact-based risk assessment would enhance use and understanding of AI; by starting from concrete risk factors, rather than abstract principles, the Framework would be approachable for "on-the-ground" software engineers, compliance officers,

[4] Exec. Order No. 13859 84 C.F.R. 3967 (2019).

[5] Framework at 86 Fed. Reg. 40811.

and corporate leadership alike. When presented with potential risks and examples of remediation strategies, stakeholders would be better equipped to realize the Framework's goals.

Explainability. Given the potential opacity issues of certain AI systems, the potential utility of risk management frameworks standards is not meaningful unless the public has some understanding of how these systems work, and the basis for their output. For this reason, CTA recommends that the Framework encourage AI developers to create new, or utilize existing, explainability tools to help users understand AI systems' decision-making processes. When affected individuals can understand the basis for a recommendation, decision or action taken by the AI system, then consumer buy-in and acceptance of this technology is likely to increase. That, in turn, can mitigate potential risks and lead to broader adoption of AI.

Diversity and Representation. A key component of trustworthiness is the minimization of potential bias in systems and their outputs. To address bias, AI developers should be encouraged to expand their use of diverse and representative data. If systems are trained on limited data sets, they are necessarily constrained by the extent of the data, which may be insufficiently representative and risk engendering bias.

II. Standards, frameworks, models, methodologies, tools, guidelines and best practices, and principles to identify, assess, prioritize, mitigate, or communicate AI risk and whether any currently meet the minimum attributes described in the Framework (RFI No. 5)

CTA recommends that the Framework development process include an assessment of existing work on this topic. Particularly, CTA recommends that NIST consider incorporating, or relying upon, several recent standards developed by CTA, including: CTA project CTA-2096 regarding developing trustworthy AI systems,[6] CTA-2089 regarding the definitions and characteristics of AI, [7] the published standard CTA-2090 regarding the use of AI in healthcare and trustworthiness,[8] and the ISO/IEC JTC1 standards project 23894 regarding AI risk management.[9] Further, the Organisation for Economic Cooperation and Development (OECD) is in the process of developing its own framework for assessing the opportunities and potential risks presented by different types of AI systems.[10]

---

[6] Consumer Technology Association, *Guidelines for Developing Trustworthy AI*, CTA-2096 (Dec. 20, 2019), https://standards.cta.tech/apps/group_public/project/details.php?project_id=637.

[7] Consumer Technology Association, *Definitions and Characteristics of Artificial Intelligence*, CTA-2089 (Mar. 4, 2020), https://standards.cta.tech/apps/group_public/project/details.php?project_id=601.

[8] Consumer Technology Association, *The Use of Artificial Intelligence in Health Care: Trustworthiness*, ANSI/CTA-2090 (Feb. 2021), https://shop.cta.tech/products/the-use-of-artificial-intelligence-in-healthcare-trustworthiness-cta-2090.

[9] International Organization for Standardization and International Electrotechnical Commission, *Artificial Intelligence Risk Management*, ISO/IEC JTC 1/SC 42.

[10] OECD Framework for the Classification of AI Systems – Public Consultation on Preliminary Findings (rel. Spring, 2021), https://aipo-api.buddyweb.fr/app/uploads/2021/06/Report-for-consultation_OECD.AI_Classification.pdf .

Additionally, existing laws and norms at the international, multi-level, national, and sub-national level address data protection and privacy, discrimination, and consumer protection. For example, AI providers in Europe are subject to certain notice, opt-out and transparency obligations under the General Data Protection Regulation (GDPR) for any data processing arising from the use of EU subjects' data. In the U.S., numerous existing federal and state laws protect against discriminatory behavior involving individuals' access to housing, finance, healthcare and other areas of fundamental rights. These laws are successful in protecting consumers, enabling recourse and setting standards and compliance expectations for companies developing AI systems.

The Framework should leverage these existing and proven laws by (a) incorporating criteria to focus on compliance with existing applicable law, and (b) emphasizing that industry groups, jurisdictions, or other actors should acknowledge that existing law already establishes important ground rules when developing self-regulatory standards or any potential new regulation.

Leveraging existing law not only ensures that the public and businesses understand their rights and responsibilities but also furthers the efficient implementation of the Framework. Businesses can utilize their existing compliance processes to ensure compliance with existing laws, resulting in further efficiency and cost reduction as they ensure their practices are aligned with the Framework and aligned with perceived and actual risk that will enhance technical trustworthiness.

III. How organizations take into account benefits and issues related to inclusiveness in AI design, development, use and evaluation—and how AI design and development may be carried out in a way that reduces or manages the risk of potential negative impact on individuals, groups, and society (RFI No. 8)

CTA highlights that organizations likely to be the most successful in trustworthy AI design and development prioritize the following AI design and development practices:

Inclusive design. System development that pursues inclusivity-by-design (as a conceptual corollary to privacy-by-design or security-by-design) has a far greater likelihood of achieving trustworthiness. Inclusivity should be stated as a guiding principle in design from the outset, such that all development stakeholders are aware of and committed to its realization.

Participatory processes. To the extent feasible, AI design and development should seek broad participation from stakeholders across an organization or field. Leaving development to engineers alone, or other related silos, risks unintentional oversight of key considerations, such as long-term risks, diversity objectives, legal or public policy perspectives, and impact on groups and individuals. A truly participatory process, bringing in the perspective of potentially affected communities, could be the gold standard of development where the potential risks of bias or discriminatory outputs are high.

Testing with diverse audiences. AI systems should be tested with diverse groups, in various geographies, and, to the extent relevant, in multiple languages. Just as testing for bugs is

standard in the software development process, so too should be testing for trustworthiness with a broad variety of audiences.[11]

IV. The appropriateness of the attributes NIST has developed for the AI Risk Management Framework (RFI No. 9)

CTA commends NIST for its thoughtful identification of attributes for the Framework. Greater specificity and precision in the attributes would strengthen the Framework and support its long-term success.

Risk versus trust. The RFI uses the terms "risk" and "trust" interchangeably. Although risk mitigation and development of trust are both necessary for the successful adoption and use of AI technology, they are different issues. It is unclear if the intention of the Framework is risk management or building trust. If the goal is risk management, there are many additional factors to consider beyond building trust.

"Risk" definition. NIST should also clarify whether the definition of "risk" includes safety-related risks as well. Many industries (such as healthcare) focus on safety-related risks, and not business risks. Risks should be quantifiable, concrete, and evidence-based, such that developers are empowered to meaningfully assess their products and remediation strategies can be identified.

Transparency in context. The Framework discusses the need for transparency, but too much transparency can also introduce new risks, such as the risk of information overload (such that users do not have a meaningful or clear understanding) and the need to protect privacy. The Framework should make clear that the goal is transparency in context, in a manner that is meaningful and clear to the user (as discussed in Section I above).

Taxonomy of "failures." Often, failures in the development of AI relate to a developer's lack of awareness about a specific piece of information, such as not knowing what questions to frame when developing the AI system, or being unaware that software or algorithms could fail in a particular way. We recommend the Framework include a taxonomy of potential "failures," enabling developers to consider critical questions like: "Did we consider selection bias?" "Did we consider data 'drift' over time?" "Did we consider the potential for users to over-trust the application?"

V. Effective ways to structure the Framework to achieve the desired goals, including, but not limited to, integrating AI risk management processes with organizational processes for developing products and services for better outcomes in terms of trustworthiness and management of AI risks (RFI No. 10)

CTA agrees with NIST's attention not only to the substance of the Framework but also to its structure and effects. Just as NIST has considered the principles of AI development, CTA suggests several principles for the Framework's structure and process.

[11] *See, e.g.*, [DWT to provide cites for this proposition].

Flexibility. The Framework must be supported in future guidance or implementation standards with attention to flexibility. An extensive, double-digit compliance process before an AI system can be developed or used may not be appropriate across AI applications and instances without regard to inherent or perceived risk, including lack of risk. Prescriptiveness may stifle innovation and use of this technology for good.

Usability. CTA commends efforts to make the Framework user-friendly and urges NIST to prioritize usability. Experimentation and consideration in the implementation of the Framework should be encouraged; the Framework can serve as a mechanism to consider the operational aspects of AI systems, but it should not form the basis for adopting strict new regulatory obligations. Otherwise, mandates that are burdensome and difficult to implement would be costly and inhibit the development and enjoyment of the benefits of AI. This would particularly harm small and medium businesses and enterprises and their customers. CTA encourages NIST to address: (a) how it recommends the Framework be implemented by policymakers in statute, regulation, or other formal policy; and (b) targeted relief or carve-outs for smaller entities, especially those with less identifiable risk or where risks are less likely to emerge unexpectedly.

Nuance. CTA suggests NIST consider the necessary nuance in developing the Framework. For instance, the Framework should recognize the benefits of broad data collection. Appropriate collection and incorporation of additional and varied data as inputs improves AI systems and helps address and mitigate harmful bias over time. Better functioning AI (i.e., AI trained on greater amounts and more diverse data, and deployed with appropriate considerations and mitigations) can provide further benefits to the public, such that the AI systems produce more accurate, trustworthy, and ethical outputs. Without sufficient and varied data inputs, desirable outputs may remain unattainable.

Iteration. CTA emphasizes that the risks and benefits of an AI system cannot be captured in a formulaic calculation. While a formula, or risk assessment framework, may be a ***starting point*** for assessing opportunities and risks, NIST should emphasize that it is only the start of a process that will require additional data, evidence, and robust risk-benefit analyses. Further iterations of the Framework and any resulting implementation recommendations should recognize these complexities. NIST might consider an incremental and iterative process for Framework development: a series of draft frameworks leading to a final framework, accounting for the evolving nature of AI technology and emerging standards.

VI. The extent to which the Framework should include governance issues, including but not limited to the make up of design and development teams, monitoring and evaluation, and grievance and redress (RFI No. 12)

As mentioned previously, CTA believes flexibility is paramount to the success of the Framework and the development of effective, trustworthy AI. Overly prescriptive principles could have significant negative impacts on innovation, particularly by small and medium businesses, ultimately hampering the benefit to the public of AI. CTA encourages NIST in its development

of voluntary governance standards and cautions that governance ***mandates*** could undermine the intention of the Framework.

To the extent governance is addressed in the Framework, CTA recommends that NIST include activities after product launch that contribute to trustworthiness, such as: monitoring product performance; gathering user feedback; where necessary and appropriate, retraining systems; and, updating products with fresh training data on regular cadences.

VII. Conclusion

The development and deployment of AI systems presents multiple challenges as well as potentially immense and transformative benefits for society. As AI systems become increasingly interwoven in all aspects of our everyday professional and personal lives, and to realize the potential to improve our lives, developers and users of technology must enhance user trust for AI systems to be accepted and widely used, and regulators must recognize the various risk environments that will control how AI systems can safely and effectively operate. As stated in Draft NISTIR 8332, "If the AI system has a high level of technical trustworthiness, and the values of the trustworthiness characteristics are perceived to be good enough for the context of use, and especially the risk inherent in that context, then the likelihood of AI user trust increases. It is this trust, based on user perceptions, that will be necessary of any human-AI collaboration."[12] As we move from research and development to actual use, the Framework must allow for maximal appropriate data collection and encourage innovation that will lead to deployment and use of trustworthy AI.

Respectfully submitted,

/s/ Douglas K. Johnson
Douglas K. Johnson
Vice President, Emerging Technology Policy
djohnson@cta.tech

/s/ Michael Petricone
Michael Petricone
Sr. Vice President, Government and Regulatory Affairs
mpetricone@cta.tech

[12] Brian Stanton & Theodore Jensen, *Trust and Artificial Intelligence*, Natl. Inst. Stand. Technol. NISTIR 8332-draft (Mar. 2021), https://doi.org/10.6028/NIST.IR.8332-draft.