epic.org

Electronic Privacy Information Center
1519 New Hampshire Avenue NW
Washington, DC 20036, USA

+1 202 483 1140
+1 202 483 1248
@EPICPrivacy
https://epic.org

FEEDBACK OF THE ELECTRONIC PRIVACY INFORMATION CENTER

to the

National Institute of Standards and Technology

Regarding the

Artificial Intelligence Risk Management Framework: Second Draft

September 28, 2022

The Electronic Privacy Information Center ("EPIC") submits the following feedback to the request for information by the National Institute of Standards and Technology ("NIST") on the second draft of the AI Risk Management Framework (hereinafter, the "AI RMF") and its corresponding Playbook.[1]

EPIC is a public interest research center in Washington, D.C., that was established in 1994 to focus public attention on emerging privacy and related human rights issues, as well as to protect privacy, the First Amendment, and other constitutional values.[2] EPIC has a long history of promoting transparency and accountability for information technology, including AI systems.[3] Over the last decade, EPIC has consistently advocated for the adoption of clear, commonsense, and actionable AI regulations.[4] EPIC has litigated cases against the U.S. Department of Justice to

[1] AI RMF Development & Request for Information, National Institute of Standards and Technology (updated Aug. 18, 2022), https://www.nist.gov/itl/ai-risk-management-framework/ai-rmf-development-request-information; AI Risk Management Framework Playbook, National Institute of Standards and Technology (Aug. 2022), https://pages.nist.gov/AIRMF/.

[2] For more on EPIC's history, see EPIC, *About EPIC* (2019), https://epic.org/about/.

[3] *See, e.g.*, *AI and Human Rights,* EPIC (2022), https://epic.org/issues/ai/; *AI and Human Rights: Criminal Legal System,* EPIC (2022), https://epic.org/issues/ai/ai-in-the-criminal-justice-system/; Comments of EPIC, *Consumer Welfare Implications Associated with the Use of Algorithmic Decision Tools, Artificial Intelligence, and Predictive Analytics*, Federal Trade Commission (Aug. 20, 2018), available at https://epic.org/apa/comments/EPIC-FTC-Algorithmic-Transparency-Aug-20-2018.pdf; Comments of EPIC, *Developing UNESCO's Internet Universality Indicators: Help UNESCO Assess and Improve the Internet*, United Nations Educational, Scientific and Cultural Organization ("UNESCO") 5-6 (Mar. 15, 2018), available at https://epic.org/apa/comments/EPIC_UNESCO_Internet_Universality_Comment.pdf.

[4] *See, e.g.*, EPIC Urges DC Council to Pass Algorithmic Discrimination Bill (Sept. 23, 2022), https://epic.org/epic-urges-dc-council-to-pass-algorithmic-discrimination-bill/; Comments of EPIC, *Intellectual Property Protection for Artificial Intelligence Innovation*, U.S. Patent and Trademark Office (Jan. 10, 2020), *available at* https://epic.org/apa/comments/EPIC-USPTO-Jan2020.pdf; Comments of EPIC, *HUD's Implementation of the Fair Housing Act's Disparate Impact Standard*, Department of Housing and Urban Development (Oct. 18, 2019),

compel production of documents regarding "evidence-based risk assessment tools"[5] and against the U.S. Department of Homeland Security to produce documents about a program purported to assess the probability that an individual will commit a crime.[6] EPIC has also regularly submitted comments urging the adoption of adequate AI regulations that meaningfully protect the public.

More recently, EPIC submitted comments to the National Security Commission on Artificial Intelligence, the U.S. Office of Science and Technology Policy, the European Commission, and the U.S. Office of Management and Budget urging the adoption of clear AI regulations that meaningfully protect individuals.[7] To establish necessary consumer safeguards in the absence of algorithmic transparency and protective AI regulations, EPIC also recently filed FTC complaints against HireVue,[8] a company that sells algorithmic employment screening tools, and AirBnB,[9] the major rental company, which claims to use opaque algorithms to evaluate the risk of potential renters and predict the likelihood of parties in Airbnb units using factors like renter age and location. EPIC has also filed a complaint with the DC Attorney General's office about five test proctoring companies[10] and a petition with the FTC for a rulemaking on commercial AI.[11] EPIC submitted comments on the first draft of the AI framework presented by NIST.[12]

---

available at https://epic.org/apa/comments/EPIC-HUD-Oct2019.pdf; Testimony of EPIC, Massachusetts Joint Committee on the Judiciary (Oct. 22, 2019), transcription available at https://epic.org/testimony/congress/EPIC-FacialRecognitionMoratorium-MA-Oct2019.pdf; Statement of EPIC, *Industries of the Future*, U.S. Senate Committee on Commerce, Science, and Transportation (Jan. 15, 2020), available at https://epic.org/testimony/congress/EPIC-SCOM-AI-Jan2020.pdf; Comments of EPIC, *Request for Information: Big Data and the Future of Privacy*, Office of Science and Technology Policy (Apr. 4, 2014), available at https://epic.org/privacy/big-data/EPIC-OSTP-Big-Data.pdf.

[5] *EPIC v. Department of Justice*, 320 F.Supp.3d 110 (D.D.C. 2018), *voluntarily dismissed*, 2020 WL 1919646 (D.C. Cir. 2020), https://epic.org/foia/doj/criminal-justice-algorithms/.

[6] *See* EPIC, *EPIC v. DHS (FAST Program)*, https://epic.org/foia/dhs/fast/.

[7] Comments of EPIC, *Solicitation of Written Comments by the National Security Commission on Artificial Intelligence*, 85 Fed. Reg. 32,055, National Security Commission on Artificial Intelligence (Sept. 30, 2020), available at https://epic.org/apa/comments/EPIC-comments-to-NSCAI-093020.pdf; Comments of EPIC, *Request for Comments on a Draft Memorandum to the Heads of Executive Departments and Agencies, "Guidance for Regulation of Artificial Intelligence Applications,"* 85 Fed. Reg. 1825, Office of Management and Budget (Mar. 13, 2020), available at https://epic.org/apa/comments/EPIC-OMB-AI-MAR2020.pdf; Comments of EPIC, *Request for Feedback in Parallel with the White Paper on Fundamental Rights*, European Commission Fundamental Rights Policy Unit (May 29, 2020), available at https://epic.org/apa/comments/EPIC-EU-Commission-AI-Comments-May2020.pdf; Comments of EPIC, *Proposal for a legal act of the European Parliament and the Council laying down requirements for Artificial Intelligence*, European Commission (Sept. 10, 2020), available at https://epic.org/apa/comments/EPIC-EU-Commission-AI-Sep2020.pdf.

[8] Complaint and Request for Investigation, Injunction, and Other Relief, *In re HireVue* (Nov. 6, 2019), https://epic.org/privacy/ftc/hirevue/EPIC_FTC_HireVue_Complaint.pdf.

[9] Complaint and Request for Investigation, Injunction, and Other Relief, *In re Airbnb* (Feb. 27, 2019), https://epic.org/privacy/ftc/airbnb/EPIC_FTC_Airbnb_Complaint_Feb2020.pdf.

[10] Complaint and Request for Investigation, Injunction, and Other Relief, *In re Respondus, Inc., et al.* (Dec. 9, 2020), https://epic.org/privacy/dccppa/online-test-proctoring/EPIC-complaint-in-re-online-test-proctoring-companies-12-09-20.pdf.

[11] *In re: Petition for Rulemaking Concerning Use of Artificial Intelligence in Commerce*, EPIC (Feb. 3, 2020), https://epic.org/privacy/ftc/ai/EPIC-FTC-AI-Petition.pdf.

[12] *EPIC Submits Feedback on NIST AI Risk Management Framework,* EPIC (Aug. 19, 2021), https://epic.org/epic-submits-feedback-on-nist-ai-risk-management-framework/.

I. EPIC Applauds NIST's Efforts to Expand the Types of AI Risks Considered by the AI RMF, including Risks to Racial Equity and Other Risks to Individuals Impacted by AI Systems, and NIST's Adoption of Broader Stakeholder Involvement in AI Development and Deployment.

EPIC commends the ongoing efforts NIST has made to incorporate substantive public input into the AI RMF development process. NIST has taken several steps to emphasize the importance of transparency and accountability of AI systems in the second draft of the AI RMF, and EPIC encourages NIST to preserve these changes as it incorporates additional feedback. In particular, EPIC recommends that NIST safeguard or further strengthen three parts the AI RMF.

First, EPIC supports NIST's decision to expand the types of AI risks considered by the AI RMF to include risks connected to racial equity, the environment, and civil society more broadly. AI actors routinely deprioritize, downplay, or ignore the risks and harms borne by those whose data is used to train AI systems and whose data is processed by those systems. Further, many AI actors fail to explicitly consider the heightened risk and impact of AI harms on already marginalized populations. The more expansive approach to AI risk adopted in the second draft of the AI RMF is in line with recommendations supported by the International Organization for Standardization ("ISO") and Organisation for Economic Co-operation and Development ("OECD"), standardizing AI guidelines and better protecting those negatively impacted by automated decision-making.[13] EPIC also strongly supports NIST's decision specifically recommending that AI actors cease development and deployment of AI systems when risks are high and harms imminent.

Second, EPIC applauds NIST's decision to include a broader set of stakeholders within the AI development process. By encouraging AI actors to consult regularly with domain experts, end users, advocacy groups, and other entities impacted by automated decision-making, AI systems become more transparent and accountable.

Third, EPIC supports NIST's decision to specifically recommend impact assessments and regular AI accuracy and bias evaluations in the second draft of the AI RMF. EPIC believes the AI RMF is most effective where it provides clear directions on what processes and procedures an AI actor should follow to minimize harm and legal uncertainty. Specific, actionable recommendations like impact assessments not only help AI actors apply NIST guidance more effectively to their own AI development processes, but also provide regulators and policymakers with clear, actionable standards that they can use to guide enforcement efforts. Recommending impact assessments and regular AI accuracy and bias evaluations is a positive step toward clarity and accountability, and EPIC encourages NIST to continue incorporating additional specific recommendations into the AI RMF.

[13] *ISO/IEC JTC 1/SC 42 Artificial Intelligence*, International Organization of Standards, https://www.iso.org/committee/6794475.html (last accessed Sept. 27, 2022); *AI Principles*, Organisation for Economic Cooperation and Development (2019), *available at* https://www.oecd.org/going-digital/ai/principles/.

While EPIC recognizes the positive changes that have been incorporated into the second draft of the AI RMF, there remain additional areas that should be improved. EPIC urges NIST to consider the following recommendations in future drafts of the AI RMF.

II. EPIC Urges NIST to (1) Include Clear Accountability Mechanisms that Meaningfully Protect Sensitive Data, (2) Discourage AI Practices like Emotion Recognition and Biometric Categorization That Have Proven and Serious Risks to Those Impacted, and (3) Incorporate Specific Examples and Recommended Processes.

EPIC Recommendation 1: Incorporate Enforcement Actions, Penalties, and Other Mechanisms to Ensure that Those Claiming to Adhere to the AI RMF Actually Do.

Without oversight, measurable demonstrations of compliance efforts, and the possibility of enforcement or other legal liability, AI actors have few incentives to spend the time and resources necessary to conform their practices to the AI RMF. NIST can and should maintain objectivity and flexibility in developing the AI RMF, but the AI RMF's usefulness as a standard framework is undercut by the ease with which AI actors can ignore NIST guidance—or falsely claim adherence without repercussion. Without actionable accountability mechanisms, there can be no public or industry trust in the AI RMF as a meaningful standard of AI development.

NIST has incorporated some evaluation and accountability procedures into the AI RMF under the Measure function, but these procedures are presented as steps that AI actors can take internally, rather than meaningful forms of accountability for AI actors that either negligently or willfully develop AI systems without considering the full range of risks involved. Allowing purely internal procedures with no transparency obligations or outside review may render these internal measures largely meaningless and creates no incentive for AI actors to expend the time and effort necessary to put them in place. Further, the lack of oversight and enforcement would allow entities to falsely claim compliance with NIST guidance, taking advantage of NIST's reputation for their own benefit while failing to enact the practices NIST seeks to promote. False claims with no review or enforcement could not only erode public trust in this AI RMF, but erode public trust in NIST as a body as well. **To ensure that the AI RMF serves as a meaningful framework that impacts the use and development of AI systems, NIST should clearly incorporate enforcement actions, penalties, or other accountability mechanisms directly within the AI RMF—and provide actionable guidance that regulators and policymakers can use to craft their own enforcement regimes.**

EPIC Recommendation 2: Implement Additional Protections for Those Impacted by AI Systems, Including Use Restrictions for Sensitive Data Categories and Restrictions on Unsubstantiated and Unsafe AI Practices like Emotion Recognition and Biometric Categorization.

For the AI RMF to be effective, it must not only highlight the practices, policies, and procedures that AI actors should follow, but also those that AI actors should avoid. Given the data-dependent nature of AI development, as well as the number of potential risks and stakeholders impacted by AI systems, risk of data misuse and potential harm is high. Even well-

meaning AI actors may unintentionally harm end users and others impacted by AI systems when developing AI if they are unaware of what AI practices are considered high-risk or have not considered disparate impact on different groups. To minimize the risk of these harms, the AI RMF must contain specific protections for sensitive and high-risk use cases. The exact form of these recommended protections may vary—and need not be exhaustive within the AI RMF—but could include data subject rights, mandatory notice to affected individuals prior to processing, consent mandates, cybersecurity measures, and explicit bans on AI systems or uses that have been proven to perpetuate discrimination. **EPIC urges NIST to incorporate the principles of data minimization, which mandates that "data should only be collected, used, or disclosed as reasonably necessary to provide the service requested by a consumer."**[14]

Specifically, EPIC encourages NIST to include additional protections for sensitive user data, including children's data and data pertaining to race, sexuality, gender, or ethnicity. AI systems that use these categories of data are particularly vulnerable to inequitable or otherwise harmful outcomes and should always be considered high-risk.[15] For example, Meta recently settled a suit brought by the U.S. Justice Department relating to its automated advertising system, which allowed landlords to target advertisements for housing options to specific populations, including according to race, gender, religion, and other sensitive characteristics.[16] The settlement mandates that Facebook overhaul its entire ad targeting tool (Lookalike Audiences) in order to ensure that housing ads reach an audience more representative of the full population.[17] In this case, the algorithm was determined to be discriminatory even though the system did not explicitly allow advertisers to click a box for "white," "black," or other groups—rather, the algorithm used proxies that *resulted* in racial discrimination, such as gender, age, interests, and zip code. This is an example of why simply stating that a system should not explicitly include a sensitive characteristic does not automatically make that system nondiscriminatory. **EPIC urges NIST to provide that algorithmic actors evaluate discriminatory *impact* of their system, beyond the use of facially discriminatory data.**

EPIC also urges NIST to explicitly denounce certain AI applications that have proven to be particularly harmful and vulnerable to algorithmic bias, including emotion recognition systems and biometric categorization systems. Emotion recognition systems rely

---

[14] *See* Consumer Reports and Electronic Privacy Information Center, *How the FTC Can Mandate Data Minimization Through a Section 5 Unfairness Rulemaking* (Jan. 26, 2022), https://epic.org/wp-content/uploads/2022/01/CR_Epic_FTCDataMinimization_012522_VF_.pdf.

[15] *See, e.g.*, Complaint for Permanent Injunction and Other Equitable Relief at 34-35, *F.T.C. v. CompuCredit Corp.*, No. 1:08–CV–1976–BBM–RGV (N.D. Ga. Oct. 8, 2008), https://www.ftc.gov/sites/default/files/documents/cases/2008/06/080610compucreditcmptsigned.pdf (FTC suit against credit card company that allegedly used undisclosed behavioral scoring algorithm to determine credit limitations based on consumer conduct); James Vincent, *The Invention of AI 'Gaydar' Could be the Start of Something Much Worse*, The Verge (Sept. 21, 2017), https://www.theverge.com/2017/9/21/16332760/ai-sexuality-gaydar-photo-physiognomy; Claudia Garcia-Rojas, *The Surveillance of Blackness: From the Trans-Atlantic Slave Trade to Contemporary Surveillance Technologies*, Truthout (Mar. 3, 2016), https://truthout.org/articles/the-surveillance-of-blackness-from-the-slave-trade-to-the-police/ (discussing Professor Simone Brown's research on how race and anti-Black colonial logics inform contemporary surveillance practices).

[16] Naomi Nix and Elizabeth Dwoskin, *Justice Department and Meta settle landmark housing discrimination case,* The Washington Post (Jun. 21, 2022), https://www.washingtonpost.com/technology/2022/06/21/facebook-doj-discriminatory-housing-ads/.

[17] *Id.*

on the false premise that both universal emotions and a clear correlation between emotion and facial expression exist—a premise that has been repeatedly disproven.[18] Similarly, biometric categorization systems are based on the belief that certain physical characteristics can be linked to specific traits. This is essentially a form of digital phrenology.[19] Companies using these system types have claimed to be able to predict anything from the likelihood of terrorist leanings to sexuality based solely on the analysis of facial features.[20] Both of these system types exhibit persistent and inherent inaccuracy and bias that cannot be divided from the systems or meaningfully resolved such that they do not cause harm. EPIC believes these systems are harmful by their very nature and urges NIST to denounce them entirely.

EPIC Recommendation 3: Add Clarity to the AI RMF by Including—Not Removing—Additional Specific Examples and Recommended Processes

As discussed above, the AI RMF is most effective where it provides clear examples of what AI actors should and should not do to minimize harm and legal uncertainty. The primary benefits to including specific examples in the AI RMF itself are clarity and expertise. AI actors, regulators, and civil society organizations know that examples presented by NIST are acceptable forms of compliance with the AI RMF. Examples give these entities clear directions in which to allocate resources and assist organizations that may not have the technical expertise to identify specific acceptable approaches on their own. Although NIST has provided opportunities for interested parties to submit use case profiles and other specific guidance through the NIST Trustworthy and Responsible AI Resource Center, these resources are not necessarily reliable guidance on AI RMF compliance. Those who author these resources are necessarily less familiar with NIST's priorities and policies than NIST itself. Additionally, the AI RMF, once finalized, is

[18] Kate Crawford, *Artificial Intelligence is Misreading Human Emotion*, The Atlantic (Apr. 27, 2021), https://www.theatlantic.com/technology/archive/2021/04/artificial-intelligence-misreading-humanemotion/618696/; Lisa Feldman Barret et al., *Emotional Expressions Reconsidered: Challenges to Inferring Emotion from Human Facial Movements*, 20 Ass'n for Psych. Sci., 1, 46 (2019), available at https://journals.sagepub.com/doi/pdf/10.1177/1529100619832930; *see also* Krys, Kuba et al., *Be Careful Where You Smile: Culture Shapes Judgments of Intelligence and Honesty of Smiling Individuals*, Journal of Nonverbal Behavior Vol. 40, 101-116 (2016), available at https://link.springer.com/article/10.1007/s10919-015-0226-4; Charlotte Gifford, *The Problem with Emotion-Detection Technology*, The New Economy (Jun. 15, 2020), https://www.theneweconomy.com/technology/the-problem-with-emotion-detection-technology.

[19] *See* Blaise Aguera y Arcas et al., *Physiognomy's New Clothes*, Medium (May 6, 2017), https://medium.com/@blaisea/physiognomys-new-clothes-f2d4b59fdd6a.

[20] *See* Sally Adee, *Controversial Software Claims to Tell Your Personality From Your Face*, New Scientist (May 27, 2016), https://www.newscientist.com/article/2090656-controversial-software-claims-to-tellpersonality-from-your-face/; *Researchers are Using Machine Learning to Screen for Autism in Children*, Duke Pratt School of Engineering (Jul. 11, 2019), https://pratt.duke.edu/about/news/amazon-autism-app-video; Paul Lewis, *"I was Shocked it was so Easy": Meet the Professor Who Says Facial Recognition Can Tell if You're Gay*, The Guardian (Jul. 7, 2018), https://www.theguardian.com/technology/2018/jul/07/artificial-intelligence-can-tell-your-sexuality-politics-surveillance-paul-lewis; Madhi Hashemi & Margaret Hall, *Criminal Tendency Detection from Facial Images and the Gender Bias Effect*, 7 J. Big Data, 1, 1 (2020), https://journalofbigdata.springeropen.com/articles/10.1186/s40537-019-0282-4#Sec9 (since retracted); Luana Pascu, *Biometric Software that Allegedly Predicts Criminals Based on Their Face Sparks Industry Controversy*, Biometric Update (May 6, 2020), https://www.biometricupdate.com/202005/biometric-software-that-allegedly-predicts-criminals-based-on-their-face-sparks-industry-controversy.

meant to be a static and reliable standard for AI risk management, while the Resource Center may be continually updated to reflect shifting guidance and priorities.

Taken together, NIST's decision to remove specific examples from the AI RMF and replace them with external resources makes it more difficult for AI actors to comply, not less. Without clear, approved examples within the AI RMF itself, AI actors are likely to either claim compliance incorrectly because they lack the technical expertise to know they are falling short or be forced to expend significant resources to consult experts on how to incorporate the AI RMF into their internal processes. Neither option benefits NIST or supports AI risk management: the former erodes public trust in NIST and the AI RMF as false claims of compliance increase, while the latter puts small and medium-sized enterprises at a disadvantage for compliance by injecting additional costs and legal uncertainty into the AI development process.

To ensure that AI actors adopt the recommendations set forth in the AI RMF, NIST should include clear, actionable recommendations on how to comply with its guidance. These recommendations might include examples of effective methods in use today, technical proposals, and case studies.

Conclusion

EPIC commends NIST's decision to expand the breadth of AI risks and stakeholders considered in the second draft of its AI RMF and recommends that NIST prioritize the following when further developing the AI RMF: (1) enforcement mechanisms to guide adherence to the AI RMF and standardize industry practices; (2) additional protections for high-risk use cases, including AI systems that use sensitive data categories and AI systems that involve emotion recognition and biometric categorization; and (3) clear examples of how AI actors can incorporate the AI RMF into their workflows. NIST can and should maintain objectivity and flexibility, but without clear, actionable guidance on what AI actors should and should not do, the AI RMF will prove difficult to meaningfully implement.

Respectfully Submitted,

/s/ *Calli Schroeder*
Calli Schroeder
EPIC Global Privacy Counsel

/s/ *Ben Winters*
Ben Winters
EPIC Counsel

/s/*Grant Fergusson*
Grant Fergusson
EPIC Equal Justice Works Fellow