hackerone

January 29, 2026

VIA ELECTRONIC SUBMISSION
CyberAIProfile@nist.gov

Re: Comments on NIST Cybersecurity Framework Profile for Artificial Intelligence (NISTIR 8596)

Dear Sir or Madam,

HackerOne Inc. (HackerOne) submits the following comments in response to the National Institute of Standards and Technology's Cybersecurity Framework Profile for Artificial Intelligence (AI).

By way of background, HackerOne is a global leader in finding and fixing critical vulnerabilities and AI security issues. Our industry-leading HackerOne Platform combines AI with the expertise of the world's largest community of security researchers to uncover and remediate vulnerabilities and AI security issues across the software development lifecycle. The platform offers bug bounty, vulnerability disclosure, pentesting, code review, and AI red teaming.

HackerOne consistently advocates for widespread adoption of cybersecurity measures that have proven effective at addressing unmitigated vulnerabilities in both commercial and government contexts. This advocacy extends to the realm of AI, where we set up bug bounties for AI security testing and help reduce undesirable outputs in AI. As the demand for secure AI grows, HackerOne is best positioned to assist entities in navigating the complexities of deploying AI models responsibly.

We recognize AI's dual role in both enabling new capabilities and introducing novel and evolving risks, and we welcome the opportunity to collaborate with NIST and other stakeholders to advance standards that promote resilience, transparency, and trust in AI systems.

In March of 2025, HackerOne submitted comments to inform the development of a Cyber AI Profile, and we appreciate that the preliminary draft reflects several of these concepts, including the integration of AI red teaming and and the recognition of vulnerability disclosure mechanisms as a component to secure AI risks. However, our prior recommendation to explicitly incorporate bug bounty programs (BBPs) as mechanisms for identifying AI-specific vulnerabilities is not reflected in the current draft. Below we provide additional context and information on why we believe BBPs should be included.

Bug bounty programs represent a mature, widely adopted security practice that incentivizes continuous, real-world testing of systems by independent security researchers operating under coordinated disclosure frameworks. In the AI context, BBPs are particularly effective at identifying vulnerabilities that may not be uncovered through other forms of testing, including prompt injection, model abuse, data leakage, unexpected model behavior, and misuse of

deployed AI capabilities. As an extension of VDPs, BBPs strengthen existing disclosure mechanisms and allow organizations to benefit from diverse perspectives that complement internal testing and red teaming efforts.

While the Cyber AI Profile appropriately emphasizes testing, assessments, and improvements derived from exercises and VDPs, the absence of explicit references to BBPs creates a gap between the Profile and current security practices used by many organizations to secure AI systems in production. Explicitly recognizing BBPs would provide clearer, more actionable guidance to organizations seeking to operationalize continuous AI security testing. Accordingly, HackerOne believes that BBPs align most directly with **CSF 2.0 Core Subcategory ID.RA-08**, which focuses on the identification and analysis of cybersecurity risks, including those surfaced through external reporting and coordinated disclosure mechanisms.

* * *

Overall, we encourage NIST to recognize the important role that both automated tools and human security researchers play in securing AI systems. HackerOne appreciates the opportunity to comment on the Cyber AI Profile and looks forward to continued engagement with NIST. We would welcome the opportunity to serve as a resource as this work continues.

Respectfully Submitted,

Ilona Cohen
Chief Legal and Policy Officer
HackerOne