Comments of the Messaging Malware Mobile Anti-Abuse Working Group (M3AAWG) on NIST AI 100-5, A Plan for Global Engagement on AI Standards

Consultation reference: NIST AI 100-5, A Plan for Global Engagement on AI Standards

The Messaging Malware Mobile Anti-Abuse Working Group (M^{3}AAWG) appreciates the opportunity to submit comments in response to the above-referenced consultation. M^{3}AAWG is a technology-neutral global industry association. As a working body, we focus on operational issues of internet abuse including technology, industry collaboration, and public policy. With more than 200 institutional members worldwide, we bring together stakeholders in the online community in a confidential yet open forum, developing best practices and cooperative approaches for fighting online abuse.

AI is a global phenomenon which impacts various countries and a number of industry sectors at high risk of abuse by cybercriminals and other threat actors. Thus, international and cross-sector engagement and involvement in standard-setting is of paramount importance. M^{3}AAWG's comments relate to three main items:

1. As a group of anti-abuse specialists, M^{3}AAWG believes that the document in its current iteration does not consider sufficiently engagement on standards regarding anti-abuse, anti-fraud, and anti-cybercrime measures.
   a. AI and ML are vulnerable to various attacks by criminals and other abusive actors. AI systems can operate at very high speed and at large scale, making inbuilt anti-abuse measures necessary as well as appropriate. Supporting anti-abuse processes is essential to mitigating social, economic, and national security risks.
   b. Therefore, we urge deep engagement with anti-abuse organizations, including, but not limited to, the Messaging Malware Mobile Anti-Abuse Working Group (M^{3}AAWG).
   c. Anti-abuse and anti-fraud are distinct from, if related to, safety, security, privacy reliability, and explainability, and should be considered in the context of the above-mentioned objectives.

2. As part of mitigating AI risks, sharing and coordination are essential. This includes but is not limited to areas such as attacks, incidents, vulnerabilities, mitigations, best practices, etc. NIST should make enabling such coordination and sharing a priority.

3. Specifically, we observe possible gaps in section 4.1 in the draft document:

a. Standards are needed to define how to demonstrate training data transparency, and to prove both authenticity and integrity of training data.
b. In the security section, standards and best practices are needed for red-teaming and pen–testing of AI systems as well as evaluating the efficacy of these procedures.
c. Standard taxonomy and terminology are needed across sectors and entities to address abuse and fraud using AI systems. NIST should work with standards organizations to develop and adopt these.

We appreciate the opportunity to submit these comments, and we welcome further opportunities to engage as needed to answer any questions during this process. Please address any inquiries to M^{3}AAWG Executive Director Amy Cadagin at comments@m3aawg.org.

Sincerely,
Amy Cadagin, Executive Director
Messaging Malware Mobile Anti-Abuse Working Group
comments@m3aawg.org
P.O. Box 9125 Brea, CA 92822