Request for Comments: AI-Generated Content Metadata Labeling Requirements

Abstract

This document outlines the regulatory framework led by the Cybersecurity and Infrastructure Security Agency (CISA) concerning mandatory metadata labeling of AI-generated content. These efforts align with global initiatives, including those by the European Union, China, and Russia, to combat disinformation and enhance digital content transparency. The implementation of these requirements begins on September 1.

1. Introduction

With the increasing integration of AI-generated content across digital platforms, governments worldwide are establishing standards to ensure transparency and accountability. CISA, in collaboration with international partners, is mandating metadata labeling for AI-generated materials to facilitate automated detection, enhance user awareness, and mitigate the spread of misinformation. This regulatory framework aligns with ongoing initiatives in China, Russia, the European Union, and the United States.

2. Scope

This RFC applies to:

- Online platforms and service providers incorporating AI-generated content.

- App store operators distributing applications that generate or manipulate content using AI.

- Users and organizations that engage with AI-generated materials and rely on metadata verification.

3. Metadata Labeling Requirements

The following metadata tagging mechanisms must be adopted by service providers:

- **Standardized Metadata Encoding:** AI-generated content must embed standardized metadata markers to allow for automated identification and verification.

- **Cryptographic Hashing for Integrity:** Metadata must include cryptographic signatures to validate content authenticity and detect unauthorized modifications.

- **Cross-Border Metadata Recognition:** Metadata labeling must be interoperable with frameworks established by China, Russia, the EU, and other jurisdictions.

- **Automated Compliance Auditing:** Platforms must implement automated systems to verify metadata compliance across digital content distributions.

4. Compliance and Enforcement

- Platforms failing to comply with metadata labeling requirements may be subject to regulatory penalties enforced by CISA and partner agencies.

- Unlabeled AI-generated content may be permitted only if it complies with designated oversight mechanisms and user disclosure policies.

- Service providers must maintain internal audit logs verifying metadata integrity and submit compliance reports upon request.

5. International Coordination

This regulatory framework is designed to align with global efforts in metadata labeling. Russia, China, and the European Union have implemented or are developing similar policies to ensure AI-generated content is identifiable and traceable via metadata tagging. CISA will coordinate with these international bodies to establish best practices and ensure cross-border metadata interoperability.

6. Implementation Timeline

- **Announcement Date:** March 2025

- **Enforcement Begins:** September 1, 2025

- **Ongoing Review:** Continuous assessment and refinement of metadata labeling protocols

7. Conclusion

The introduction of metadata labeling requirements for AI-generated content represents a significant advancement in ensuring transparency, integrity, and accountability in digital media. By embedding identifiable metadata in AI-generated materials, stakeholders can effectively detect, verify, and track content provenance. These measures will enable a more secure and verifiable information ecosystem while maintaining alignment with international regulatory efforts.

8. References

- Cybersecurity and Infrastructure Security Agency (CISA) Official Guidance

- Cyberspace Administration of China (CAC) Metadata Labeling Regulations

- European Union AI Regulation Framework

- Russian Federation Digital Content Transparency Guidelines

- United States AI Metadata Labeling Policies