OpenAI

September 5, 2025

Outbound coordinated disclosure policy

Purpose

OpenAI is committed to enhancing global software security by responsibly disclosing vulnerabilities discovered in third-party software. This policy governs how OpenAI reports and discloses vulnerabilities we discover in third-party software to vendors and open source maintainers. In these reports we aim to be cooperative, impactful, and maintain high integrity standards.

Scope

This policy lays out how we disclose issues found in open-source and commercial software through automated and manual code review, as well as discoveries arising from internal usage of third-party software and systems. Detection methods may include:

- AI- or agent-powered application security analysis.
- Security research, audits, and fuzzing of open-source software.
- Evaluations of vendor-supplied or third-party software used within OpenAI operations.
- Other methods as appropriate.

OpenAI

Principles

Ask ChatGPT

Ecosystem security: A primary goal is to improve ecosystem security.

Cooperative: Good faith, helpful engagements.

Discrete by default: Initial disclosures are private. Public disclosures occur only after explicit vendor consent, active exploitation, or if necessary to protect OpenAI and its products or customers, or the public interest.

High scale, low friction: Validated, actionable disclosures are sent. Third-party inbound disclosure processes are respected, where possible.

Attribution when relevant: Vulnerabilities will be credited as discovered by OpenAI. Where applicable, we may attribute discovery to specific individuals, systems, or agents.

Disclosure workflow

1. **Identification and Validation**
   Upon discovery, each finding is validated for security impact. Reports typically include:
   - Impact summary (e.g., denial-of-service, memory corruption)
   - Affected versions / commit ranges
   - Reproduction steps or a proof-of-concept (PoC) where possible
   - Reproduction aids (e.g., Docker image) where feasible
2. **Peer Review**
   Each disclosure undergoes an internal review process prior to release, including peer review by a security engineer for accuracy, reproducibility, and quality.
   - Where a vulnerability is discovered by an automated system, a security engineer reviews the disclosure before it is released.

A security program manager coordinates disclosures, maintains records, and manages vendor interactions.

3. **Disclosure Process**
   We will generally seek to follow the inbound disclosure procedures of the software or system being disclosed on. While we offer this in good faith, we reserve the right to follow an alternate procedure at our discretion. Our preferred methods for disclosure include vendor security emails (e.g. the company's Product Security Incident Response Team email intake), CERTs (CERT/CC, CISA), or private GitHub reporting. Submissions to public trackers such as GitHub Issues are avoided by default. We generally will not participate in Bug Bounty programs.OpenAI's outbound disclosure communications are managed through a mailing list dedicated for this purpose. Access to this mailing list will be internally restricted in the interest of confidentiality and discretion.Internally, a member of our disclosure team will examine our own use of the affected product or code. If use is found, we engage appropriate teams in accordance with our policies.
4. **Public Disclosure Timeline**

We do not commit to strict publication timelines. Public disclosures occur when one of the following occurs:

- Explicit vendor permission is obtained.
- There is credible evidence of active exploitation.
- Disclosure is essential for protecting OpenAI, its users, or the public interest.

There may be exceptions in certain cases. For example, if we report an issue to a vendor and do not receive a response, we may elect to publicly disclose the issue. Additionally, if timelines shift or there's a material difference in views around severity or patching, we may proceed with disclosure to ensure transparency and consistency in our process.

Looking forward

implementing it. We anticipate that these will change substantially as security research and application security testing become increasingly automated with advances in AI.

OpenAI

Developer Forum ↗

Manage Cookies

English United States