

EU AI Act

Compliance assessment — 2024/1689

Server: Cloaked Agent

Slug: cloaked-agent

Scan id: 1d632a68-d297-4906-bbc0-0c7282ac3d16

Assessed at: 2026-05-11 09:30:23 UTC

Sentinel version: 0.4.0

Rules version: 2026-04-23

Non-compliant

DRAFT for review — not legal advice. See attestation block for verification instructions.

Table of contents

1. Executive summary
2. Coverage & transparency
3. Controls summary
4. Control details
5. Multi-step attack chains
6. Cryptographic attestation

1. Executive summary

Assessment of Cloaked Agent against EU AI Act: overall status non compliant. Of 5 controls, 3 met, 1 unmet, 1 partial, 0 not applicable. 5 control(s) fell within MCP Sentinel's current assessor coverage; remaining control(s) are documented as not_applicable until Phase 6 expands coverage. Unmet controls have findings at or above the framework's mandatory severity threshold and should be remediated before relying on this server in a regulated deployment. All claims are traceable to individual finding rows via finding_id and to the governing rule via rule_id; the enclosing signed envelope commits MCP Sentinel to the exact bytes of this report.

2. Coverage & transparency

Coverage band: low

Coverage ratio: 40%

Rules version: 2026-04-23

Analysis techniques applied:

- ast-taint
- capability-graph
- entropy
- linguistic-scoring
- schema-inference

3. Controls summary

ControlName

Status

Evidence

Art.9Risk

Management

System' Unmet

1

Art.12Record-

Keeping' Me@

Art.13

Transparency
& Provision of
Information to
Deployers' Met
0

Art.14 Human
Oversight' Met
0

Art.15
Accuracy,
Robustness,
and
Cybersecurity!
Partial1

4. Control details

Art.9 — Risk Management System

' Unmet

22 assessor rule(s) evaluated this control; 1 finding(s) observed (1 high); at least one finding is at or above the high threshold (status: unmet).

Evidence:

[High] D3
(finding
4327a933-
f46a-4e1f-bad0-
fb11ff7880ab,
confidence 88%)

SOURCE: external-content at npm:chai@4 — Dependency names are external content resolved from public package registries. A near-miss to a popular canonical name is a supply-chain anomaly under ISO 2700

Required mitigations:

- Verify that the flagged dependency is the package you intended to install. Open the registry page for the candidate and compare publisher, publish date, download count, and postinstall scripts against the target. If it is not the intended package, replace it with the legitimate target, regenerate the lockfile, and audit the install environment (CI and developer machines) for any artifacts the malicious package may have written. Adopt a typosquat-aware package firewall (Socket.dev, Snyk Advisor, GitHub Dependabot) that rejects near-miss names at install time, in line with ISO 27001 A.5.21 supply-chain controls.

Art.12 — Record-Keeping

' Met

5 assessor rule(s) evaluated this control; no findings observed.

Art.13 — Transparency & Provision of Information to Deployers

' Met

14 assessor rule(s) evaluated this control; no findings observed.

Art.14 — Human Oversight

' Met

13 assessor rule(s) evaluated this control; no findings observed.

Art.15 — Accuracy, Robustness, and Cybersecurity

! Partial

111 assessor rule(s) evaluated this control; 1 finding(s) observed (1 medium); all findings are below the high threshold (status: partial).

Evidence:

[Medium] E1
(finding
423c8133-a9aa-
4e95-995e-3db7
e4b9f86a,
confidence 75%)

SOURCE: environment at capability:tools — An MCP server that answers tool enumeration without authentication trusts the network. Under modern threat models (CCS 2007 DNS rebinding, open cloud networki

Required mitigations:

- Require authentication for all MCP server connections. For remote MCP servers adopt OAuth 2.0 per RFC 9700 / the MCP Authorization specification. For stdio-launched servers rely on the parent process's security boundary and DO NOT expose the same server over network transports. Even localhost-bound servers should require auth: DNS rebinding (CCS 2007) makes localhost reachable from any browser tab.

5. Multi-step attack chains

No multi-step attack chains were synthesized for this server.

6. Cryptographic attestation

Algorithm: HMAC-
SHA256

Key ID: mcp-sentinel-
dev

Signer: mcp-sentinel/
v1

Signed at: 2026-05-14
T09:28:42.520Z

Canonicalization:
RFC8785

HMAC-SHA256 signature (base64, wrapped at 64 chars):

h5rFnKW4gRzc2VfJ+V3K0gKracZmGJ2x5AOkSmVr78k=

Verification instructions:

To verify this report:

1. Extract the report body (everything except the .attestation field).
2. Canonicalize the body via RFC 8785 (JCS).
3. Compute HMAC-SHA256 with the signing key for key_id "mcp-sentinel-dev".
4. Base64-encode the result and compare with the signature above.

