

Loopjacking evidence and reproduction archive

PRIMARY

Publisher	GitHub
Published	2026-09-10
Retrieved	2026-09-24
Cite key	github2026loopjacki
URL	https://github.com/adithyan-ak/loopjacking

Evidence and reproduction archive for the Loopjacking research. Four experiment harnesses with their own instructions; verification via a Python 3.10+ script checking SHA-256 manifests; frozen package versions and dependency pins.

Table: Agno AgentOS - post-approval state substitution - 2.5.6 to 3.0.9. LangGraph Agent Server - conditional post-approval substitution - 0.7.5 to 0.14.0 (0.7.4 lacks the feature). OpenClaw - representation mismatch - 2026.2.23 vulnerable, 2026.2.24 patched. OpenAI Agents SDK - negative control - 0.22.0 and 0.22.2 reject mutation.

Described as 'a purposive comparative archive, not a prevalence survey'. Neither Agno nor LangGraph has a vendor-fixed release in this archive. Evidence cutoff 10 September 2026.

Note on a discrepancy: the repository table labels the version below each affected range as safe, while the paper's own text is narrower - 2.5.5 executed the continuation and allowed direct unapproved B, which is why it was not a qualifying Loopjacking result. The paper's wording is the one to rely on.