Warum scheitern DLP, KI-Firewalls und private LLMs daran, KI-Datenlecks zu stoppen?

Kurze Antwort: Jeder Ansatz adressiert nur einen Teil des Problems – und das auf der falschen Ebene. DLP kann umgangen werden, KI-Firewalls sitzen am Netzwerkrand und sind blind für Aktivitäten auf dem Gerät, und private LLMs sind teuer, verhindern aber trotzdem nicht, dass Mitarbeiter öffentliche Tools nutzen.

Wo die einzelnen Ansätze zu kurz greifen

- **Private LLMs** sind teuer im Aufbau und Betrieb und verhindern trotzdem nicht, dass jemand ChatGPT in einem Browser-Tab öffnet.
- **KI-Firewalls** prüfen den Netzwerkverkehr, wodurch sie blind sind für das, was auf dem Endpunkt selbst geschieht, sowie für Insider-Aktivitäten.
- **Klassisches DLP** wurde für Dateien und E-Mail-Flüsse entwickelt und wird routinemäßig über Wege umgangen, die es nie überwachen sollte.

Das umgekehrte Logikproblem

Mehrere Ansätze in dieser Kategorie versuchen, Daten vor KI zu schützen, indem sie *eine weitere* KI einsetzen, um die Ausgabe zu filtern. Das erzeugt zusätzliche Kosten und eine neue Abhängigkeit, ohne dort anzusetzen, wo die Daten tatsächlich abfließen – am Endpunkt, im Moment des Zugriffs.

FortressAI setzt auf Kernel-Ebene an, die unterhalb all dieser Schichten liegt, und greift unabhängig davon, welches Tool, welcher Browser oder welche Erweiterung im Spiel ist.

Revision #1
Created 2026-07-24 04:36:21 UTC by Dennis Underwood
Updated 2026-07-24 04:36:21 UTC by Dennis Underwood