Por que a DLP, os firewalls de IA e os LLMs privados não conseguem impedir vazamentos de dados de IA?

Resposta curta: Cada abordagem trata de parte do problema a partir da camada errada. A DLP pode ser contornada, os firewalls de IA ficam na borda da rede e são cegos à atividade no dispositivo, e os LLMs privados são caros, sem fazer nada em relação aos funcionários que usam ferramentas públicas de qualquer forma.

Onde cada abordagem falha

- Os **LLMs privados** são caros de construir e operar, e não impedem que ninguém abra o ChatGPT em uma aba do navegador, independentemente disso.
- Os **firewalls de IA** inspecionam o tráfego de rede, o que os deixa cegos ao que acontece no próprio endpoint e à atividade interna (insider).
- A **DLP tradicional** foi projetada para arquivos e fluxos de e-mail, e é rotineiramente contornada por caminhos que nunca foi construída para monitorar.

O problema da lógica invertida

Várias abordagens nessa categoria tentam proteger os dados da IA inserindo *outra* IA para filtrar a saída. Isso adiciona custo e uma nova dependência sem tratar o ponto onde os dados realmente saem — o endpoint, no momento do acesso.

O FortressAI aplica a proteção no kernel, que está abaixo de todas essas camadas e se aplica independentemente de qual ferramenta, navegador ou extensão esteja envolvida.

Revision #1
Created 2026-07-24 05:48:26 UTC by Dennis Underwood
Updated 2026-07-24 05:48:26 UTC by Dennis Underwood