Why do DLP, AI firewalls, and private LLMs fail to stop AI data leaks?

Short answer: Each addresses part of the problem from the wrong layer. DLP can be bypassed, AI firewalls sit at the network edge and are blind to activity on the device, and private LLMs are expensive while doing nothing about employees using public tools anyway.

Where each approach falls short

- **Private LLMs** are costly to build and run, and they don't prevent anyone from opening ChatGPT in a browser tab regardless.
- **AI firewalls** inspect network traffic, which leaves them blind to what happens on the endpoint itself and to insider activity.
- **Traditional DLP** was designed for files and email flows, and is routinely bypassed by paths it was never built to watch.

The upside-down logic problem

Several approaches in this category try to protect data from AI by inserting *another* AI to filter the output. That adds cost and a new dependency without addressing where the data actually leaves — the endpoint, at the moment of access.

FortressAI enforces at the kernel, which is below all of these layers and applies no matter which tool, browser, or extension is involved.

Revision #3
Created 2026-07-20 19:23:58 UTC by Dennis Underwood
Updated 2026-07-21 19:32:13 UTC by Dennis Underwood