The Governance Gap No One Has Solved:

Shadow AI on Personal Devices

Author: Mark Austin, Cetus AI

Published: July 2026

Classification: Public

> *"You can't build a wall around personal devices. But you can build a better house inside the wall."*

Every organisation with an AI governance strategy has the same blind spot. You can enforce policies on corporate laptops. You can route traffic through proxies on managed networks. You can block AI services at the firewall. But the moment an employee picks up their personal phone, connects to their home WiFi, and types a project name into ChatGPT — your entire governance framework becomes irrelevant.

IBM's Cost of a Data Breach Report 2025 found that one in five breaches now involve shadow AI, adding an average of $670,000 to each incident [1]. Menlo Security's telemetry data reveals that 68% of employees use free-tier AI tools via personal accounts, with 57% inputting sensitive corporate data [2]. Zscaler recorded more than 410 million data loss prevention policy violations from ChatGPT alone in 2025 [3]. These are not hypothetical risks. They are last year's numbers.

The uncomfortable truth is that every vendor solution on the market today requires one of two things: a managed device, or traffic flowing through a corporate network. Neither condition exists when someone is on a personal phone, on personal data, asking an AI to help them draft a board paper.

Why This Is Harder Than Shadow IT

The comparison to shadow IT is instructive but misleading. When employees adopted Dropbox and Slack without approval a decade ago, the data at least sat in a definable location. Security teams could discover the app, assess the risk, and either sanction it or block it. The file stayed in one place.

Shadow AI is fundamentally different. The data does not sit anywhere — it transits. A developer pastes production logs into a chatbot to debug an issue. A geologist uploads exploration results to get a summary. A finance analyst feeds quarterly forecasts into a writing assistant. The data leaves the organisation through behaviour that looks like

normal work, produces no file transfer alert, and generates no incident ticket. By the time anyone notices, the information has been processed by a third-party model provider that sits entirely outside corporate control.

The scale is staggering. Enterprises transferred more than 18,000 terabytes of data to AI applications in 2025 — a 93% increase year-over-year [3]. Teramind's Shadow AI Report 2026 found that 89% of workplace AI use escapes enterprise governance entirely [4]. And Gartner projects that more than 40% of enterprises will experience a shadow AI-related data breach by 2030 [5].

The Personal Device Problem: An Honest Assessment

Let me be direct about what the current market can and cannot do.

Approach	Managed Devices	Corporate Network	Personal Device + Personal Data
Endpoint DLP (Purview, Forcepoint)	Yes	N/A	No
CASB / SWG (Zscaler, Netskope)	Yes	Yes	No
Network-level blocking	N/A	Yes	No
Browser isolation (Menlo, Island)	Yes (managed browser)	Yes	No
MDM / MAM	Yes (enrolled devices)	N/A	No
Policy and training	Partial	Partial	Partial
Sanctioned AI tools (the "carrot")	Partial	Partial	Partial

Every technical control in the market today has the same limitation: it requires either ownership of the device or ownership of the network path. When neither condition is met

— which is the case for every employee with a smartphone and a mobile data plan — the governance framework has a hole in it.

This is not a failure of any vendor. It is a structural constraint of the problem. You cannot inspect traffic you never see. You cannot enforce policy on a device you do not manage. And you cannot prevent someone from typing sensitive information into a browser on their own phone.

What Actually Works (And What Doesn't)

Having spoken with security leaders across mining, financial services, and government, the organisations making progress on this problem are not trying to build a wall around personal devices. They are pursuing a combination of strategies that reduce the incentive, raise the awareness, and limit the blast radius.

Strategy 1: Make governed AI better than the alternative

The single most effective control is not a control at all — it is a better product. Organisations that provide enterprise AI tools that are genuinely faster, better integrated, and less restrictive than the free alternatives see significantly lower shadow AI usage. The logic is simple: if the governed tool is more convenient than the ungoverned one, most employees will use it. This means zero-friction deployment (browser extensions, mobile apps, Teams integration), no artificial usage caps, and access to the same frontier models employees would otherwise use on personal accounts.

Strategy 2: Classify data, not devices

Accept that you cannot control every device. Instead, control where sensitive data lives and how it can be accessed. If an employee cannot export a board paper to their personal device in the first place, they cannot paste it into ChatGPT on their phone. This means

tightening data egress from corporate systems — restricting copy, download, and screen-capture from sensitive repositories — rather than trying to govern the device itself.

Strategy 3: Identity-based governance at the data source

If you control the systems where sensitive data originates — SharePoint, SAP, your ERP, your geological database — you can enforce that any AI interaction with that data flows through a governed channel. The governance layer sits between the data and the AI, regardless of which device initiated the request. This does not prevent someone from typing something from memory into a personal chatbot, but it prevents bulk extraction and systematic leakage.

Strategy 4: Detection over prevention for the residual risk

For the gap that remains — an employee recalling sensitive information and typing it into a personal AI tool — the honest answer is that prevention is not technically feasible without owning the device. What is feasible is detection: canary tokens embedded in sensitive documents, watermarking of confidential content, and monitoring for indicators that proprietary information has surfaced in unexpected places. This shifts the model from "prevent all leakage" to "detect and respond to material leakage" — which is the same model organisations already use for insider threat.

Strategy 5: Cultural and contractual reinforcement

Clear acceptable use policies, regular training that explains why (not just what), and contractual obligations around data handling create a deterrent effect. They will not stop deliberate exfiltration, but they address the 90% of shadow AI usage that is well-intentioned productivity behaviour. When employees understand that pasting a client's financial data into ChatGPT creates a genuine legal and commercial risk — not just a theoretical one — most will stop doing it.

Where the Market Is Heading

Several emerging approaches show promise for closing the personal device gap, though none are mature today.

Browser-based identity layers

Rather than managing the device, manage the identity. If an employee authenticates to access corporate data through a browser-based workspace, governance can be applied at the session level regardless of the underlying device. Vendors like Island (enterprise browser) and Venn (virtual workspace for BYOD) are pursuing this model. The limitation is adoption — it requires employees to use a specific browser or workspace for work activities.

AI-aware identity governance

The most promising architectural direction treats AI access like any other privileged access. If an employee wants to use AI with corporate data, they authenticate through a governed gateway that applies policy, inspects content, and maintains an audit trail. The AI interaction happens through the gateway regardless of device. This is the model that Songlines Control implements for managed AI — extending it to cover the "personal device reaching corporate data" scenario is the logical next step.

Behavioural analytics and anomaly detection

Rather than trying to prevent every interaction, monitor patterns that indicate systematic data extraction. An employee who suddenly starts copying large volumes of text from internal systems may be feeding it to an external AI tool. This does not require device management — it requires monitoring at the data source.

Regulatory pressure

The EU AI Act, Australia's proposed mandatory guardrails, and sector-specific regulations (APRA CPS 230 for financial services, PSPF for government) are all creating legal obligations around AI data handling. As these mature, the "we can't control personal devices" argument becomes less acceptable to regulators. Organisations will need to demonstrate they have taken reasonable steps — even if those steps cannot achieve 100% prevention.

A Practical Framework for Today

For organisations grappling with this problem right now, the following framework acknowledges the technical limitations while building a defensible position.

Layer	Action	Addresses
Provide	Deploy enterprise AI tools that are better than free alternatives	Reduces incentive to use personal tools
Protect	Tighten data egress from corporate systems (restrict copy/export of sensitive data)	Limits what can reach personal devices
Govern	Route all AI interactions with corporate data through a governed gateway	Ensures policy enforcement where data originates
Detect	Monitor for indicators of data leakage (canary tokens, behavioural analytics)	Identifies material leakage after the fact
Educate	Regular, specific training on why shadow AI creates risk (not just policy compliance)	Addresses well-intentioned misuse
Contract	Update employment contracts, supplier	Creates legal deterrent and obligation

	agreements, and JV terms to address AI data handling	

This is not a perfect solution. A determined employee on a personal device can still type sensitive information into a personal AI tool, and no technical control will prevent it. But this framework reduces the probability, limits the blast radius, and creates a defensible governance position — which is what regulators, boards, and joint venture partners require.

The Honest Conclusion

Nobody has fully solved shadow AI on personal devices. Any vendor claiming otherwise is either overstating their capability or redefining the problem to exclude the hardest part.

What organisations can do is make the governed path easier than the ungoverned one, control data at the source rather than at the device, detect material leakage when it occurs, and create cultural and contractual expectations that address the human behaviour driving the risk.

The 63% of organisations with no AI governance policy at all [1] have a more immediate problem than personal devices — they have no governance anywhere. For the organisations that have built governance for their managed environment, the personal device gap is the next frontier. It will not be closed by a single product or a single policy. It will be closed by a layered approach that accepts imperfection and optimises for risk reduction rather than risk elimination.

The question is not "can we stop every employee from using AI on their personal phone?" The answer to that is no. The question is "have we taken reasonable, proportionate steps to protect our most sensitive data from uncontrolled AI exposure?" That is a question every organisation can answer affirmatively — if they are honest about what works and what does not.

References

[1] IBM, Cost of a Data Breach Report 2025, https://www.ibm.com/reports/data-breach

[2] Menlo Security, 2025 Report: How AI is Shaping the Modern Workspace, August 2025, https://www.menlosecurity.com/press-releases/menlo-securitys-2025-report-uncovers-68-surge-in-shadow-generative-ai-usage-in-the-modern-enterprise

[3] Zscaler ThreatLabz, 2026 AI Security Report, April 2026, https://www.zscaler.com/blogs/product-insights/shadow-ai-data-risk-30-day-containment-strategy

[4] Teramind, Shadow AI Report 2026, https://www.teramind.co/l/shadow-ai-report-2026/

[5] Gartner, cited in Blackfog, The Rise of Shadow AI, February 2026, https://www.blackfog.com/the-rise-of-shadow-ai-adx-in-the-age-of-chatgpt/

Cetus AI Pty Ltd | ABN 16 695 570 819

hello@cetusai.com.au | cetusai.com.au