

# **Death of the Dark Room: How Generative AI Broke Enterprise IT's Political Cover**

*Two simultaneous fears – visible failure and visible success – permanently reshape enterprise AI governance.*

**Krzysztof Dyki**

Quasen

ORCID: [0009-0007-4511-799X](https://orcid.org/0009-0007-4511-799X)

March 2026

Strategic Briefing

*This work is licensed under a Creative Commons Attribution-NoDerivatives 4.0 International License (CC-BY-ND 4.0).  
You may share and redistribute this material in any medium or format, provided you give appropriate credit to the author  
and do not modify or create derivative works.*

---

# Death of the Dark Room: How Generative AI Broke Enterprise IT's Political Cover

Krzysztof Dyki

Quasen

ORCID: [0009-0007-4511-799X](https://orcid.org/0009-0007-4511-799X)

March 2026

---

## Abstract

Enterprise IT has historically operated inside a “dark room” – a zone of technical opacity where project failures could be rationalised through complexity narratives inaccessible to non-technical leadership. This opacity was not a flaw but a governance equilibrium: boards accepted it because the cost of independent monitoring exceeded the perceived benefit. This strategic briefing argues that generative AI has structurally destroyed this equilibrium through what the author terms the **consumer benchmark effect**: every executive now carries a consumer-grade AI tool that functions as a live performance benchmark against enterprise IT delivery, collapsing monitoring costs to near zero. The result is not simply increased accountability but the emergence of two simultaneous boardroom fears – fear of visible failure and fear of visible success. Unlike earlier consumerisation waves, which raised capability expectations, generative AI changes evaluative capacity: non-specialists can now directly judge the quality of knowledge-work outputs. Because this shift is structural, not cyclical, the standard of explainability boards will accept has been durably raised. The implications for CIO mandate, governance architecture, and enterprise AI adoption sequencing are lasting.

**Keywords:** death of the dark room, consumer benchmark effect, controlled mediocrity trap, enterprise AI transparency, AI governance, boardroom AI risk, IT accountability, organizational resistance to AI, shadow AI, CIO strategy, generative AI adoption barriers, information asymmetry collapse

**JEL Classification:** M15 (IT Management) · O33 (Technological Change: Choices and Consequences) · G34 (Mergers; Acquisitions; Restructuring; Corporate Governance) · D82 (Asymmetric and Private Information; Mechanism Design)

For thirty years, a failed ERP implementation could be buried inside an architecture diagram – a failed AI chatbot gets screenshotted and forwarded to the board before lunch.

For decades, enterprise IT could fail in silence. Not because the failures were invisible in some absolute sense. Standish, Gartner, McKinsey, auditors, and angry programme sponsors all existed. But between the board and the wreckage stood a buffer of technical language, vendor diagrams, and procedural explanation. Generative AI did not merely make IT more visible. It destroyed the political cover that technical opacity had provided.

## The dark room

The dark room was not a flaw. It was a governance equilibrium.

The logic follows the principal-agent chain formalised by Jensen and Meckling and extended to professional knowledge contexts by Sharma. Boards delegated technology decisions to specialists. Specialists controlled the language, the pace, and the criteria of evaluation. Information about failure was available, but the cost of attending to it remained too high for non-technical leadership to bear. More importantly, boards often lacked not just information but evaluative competence. They could inspect spend, deadlines, and vendor contracts. They could not readily judge whether a system was good.

That distinction matters. The asymmetry was attentional as much as informational. A failed ERP rollout could always become public, and some did. But those cases usually required whistleblowers, journalists, auditors, or post-mortem litigation. Day-to-day underperformance remained inside the room. A system could be late, over budget, awkward, and politically protected at the same time.

This dark room was never universal. Tech companies, start-ups, and genuinely digital firms did not live in it to the same degree. Nor was it necessarily malicious. In many organisations it served a legitimate function: it gave technology teams room to experiment, fail, and recover without premature board intervention. Even in regulated sectors, however, formal oversight focused on controls, infrastructure, and financial exposure. It did not give boards a practical way to judge the quality of knowledge-work outputs.

Board governance has improved since the era when Nolan and McFarlan could say most boards remained in the dark about IT. But the improvement has been strongest in process oversight, not output evaluation. McKinsey reported in late 2025 that only about 15% of boards currently receive AI-related metrics, while ISS-Corporate found that AI oversight disclosure among S&P 500 companies rose more than 84% year on year in 2024. The disclosure data tracks attention, not competence. That is exactly the point: boards are now looking at AI before they know how to judge it.

That is the model now ending.

## The consumer benchmark effect

What killed the dark room was not simply that AI is visible. It was that every executive now carries a benchmark in their pocket.

Earlier consumerisation waves changed capability expectations. The iPhone and Dropbox created a simple complaint: *I can do this at home, why can't I do it at work?* Generative AI goes further. It changes evaluative capacity. A director who cannot read an architecture diagram can still form a direct judgement about a summary, a search result, a recommendation, a draft, or a chatbot answer.

The comparison is no longer *vendor promise versus enterprise delivery*. It is *my lived experience versus your expensive system*.

That is the consumer benchmark effect. It collapses the cost of comparison to near zero. EPAM observed in 2025 that many executives were already benchmarking enterprise AI progress against tools such as ChatGPT or GitHub Copilot. McKinsey identified a parallel collapse in cost opacity through FinOps and consumption metering. The two mechanisms are different, but they converge on the same destination: FinOps collapses cost opacity; consumer AI collapses performance opacity. Together, they close the last opacity gap.

The benchmark is not scientifically clean. Sean Falconer is right to argue that consumer AI and enterprise AI often inhabit different problem spaces. A board member testing a public model on general knowledge has not proved that a proprietary legal RAG system is weak. Hilb is also right that asymmetry may shift from data access to analytical capability. The consumer benchmark confers evaluative confidence, not evaluative competence.

That objection misses the political point. Governance action is triggered by confidence, not by methodological purity. The board does not need a technically valid comparison to demand an explanation. It only needs a comparison that feels legible. The burden of proof shifts from “prove this doesn’t work” to “explain why your \$5M system feels less capable than my \$20 subscription”.

This is the panopticon dynamic. Even when the board is not actively benchmarking, IT leadership behaves as if it could. The power lies in the possibility of comparison. Once any executive can independently test AI quality, the old assumption of deferred judgement disappears.

## **Two fears, one boardroom**

### ***Fear of visible failure***

The first fear is the familiar one. AI fails in public.

Hallucinations become screenshots. A bad answer becomes a board packet. A clumsy deployment becomes reputational risk, compliance exposure, or litigation risk. This fear is not new in kind. Enterprise technology has always carried failure risk. What is new is the calibration point. Failure is no longer measured only against contractual scope, project milestones, or vendor assurances. It is measured against an ambient consumer benchmark that every board member can experience directly.

This is one reason the current deployment statistics matter. S&P Global reported in 2025 that the share of companies abandoning most of their AI initiatives had risen from 17% to 42%, with organisations scrapping an average of 46% of proof-of-concept projects before production. BCG found that only 5% of companies qualified as “AI future-built”, while 60% reported minimal gains. MIT’s NANDA project added a weaker but still useful warning signal: in preliminary findings based on 52 interviews, 153 surveys, and more than 300 disclosed initiatives, 95% of organisations reported zero financial return from generative AI investments.

These figures do not measure the same thing, and they should not be conflated. S&P tracks abandonment, BCG tracks maturity, and MIT tracks financial return. Nor do they prove that politics is the sole cause. Technical difficulty is real. But the death of the dark room compounds those technical challenges by changing the cost of failure. Organisations that once learned in private are now judged in public before they have learned enough.

The rational response is caution. Destroying IT’s protective cover does not just increase accountability. It disincentivises risk-taking. When every failure is instantly legible, freezing adoption can look more prudent than learning fast.

## ***Fear of visible success***

The second fear is less discussed and more dangerous. AI works in public.

Pfeffer and Salancik gave the underlying logic decades ago: power accrues to those who control critical uncertainties. In large organisations, complexity is not just a technical condition. It is also a political asset. The person who can explain the maze often acquires power from the maze.

Visible AI success threatens that arrangement. A system that genuinely compresses legal review time, reduces service headcount, automates reporting, or removes managerial layers does not merely create efficiency. It makes existing jurisdictions contestable. Claims to indispensability weaken. Routines that once looked necessary begin to look ornamental.

This is where the institutional immune response appears. Birkinshaw and Ridderstråle described the corporate immune system long before the current AI cycle. Ben-Zur captured the same reflex in the AI context. Anyone who has spent time inside a large institution knows its symptoms: *it won't be compatible with our systems; it can't be done here; it will take years; it will cost a fortune; it is probably illegal; who takes responsibility?* Some of these objections are legitimate. That is precisely why the immune response is effective. It rarely arrives dressed as panic. It arrives dressed as process.

The important extension is this: the resistance appears to intensify as AI efficacy becomes more visible. We can ground the mechanism theoretically, but we cannot yet prove the scaling relationship conclusively. Still, the available evidence points in the same direction. Deloitte and HKU identify fear of replacement and self-image as core barriers to AI readiness. HBR's 2026 data on "AI angst" suggests that the people using AI most intensely can also be the most anxious about it. The better the system works, the sharper the jurisdictional threat.

These attributions are primary, not exclusive. Boards can also fear success, especially if it exposes their earlier inattention or forces restructuring they are not ready to govern. Managers can also fear failure. But the centres of gravity are different: boards principally fear visible failure; organisational agents principally fear visible success. The CIO sits where the two lines cross.

## ***The controlled mediocrity trap***

Once those two fears are present at the same time, the rational governance posture is neither bold adoption nor outright refusal. It is controlled mediocrity.

If the CIO must simultaneously manage the risk of visible failure and the risk of visible success, the safest option is to deploy AI that is good enough to avoid embarrassment but not so effective that it threatens established structures. That is the controlled mediocrity trap.

This is why the best sentence in many enterprise AI meetings is also the most dangerous one: "It works, but only within a tightly scoped use case". Sometimes that is disciplined execution. Often it is political self-defence.

What the consumer benchmark frames as mediocrity, the enterprise CIO may rightly call appropriate scoping. Security, traceability, integration, and compliance are real constraints. But the political distinction dissolves when the board's reference point costs 0.4% of the enterprise budget. Dataiku's 2026 CIO survey captured the mood precisely: 74% of CIOs regretted at least one major AI vendor or platform decision, 71% expected AI budgets to be cut or frozen if targets were missed, and 29% said they had already been asked multiple times to justify outcomes they could not fully explain.

*The 5% who achieve value at scale are not buying better AI. They are building organisations that can absorb it.*

## The one-way door

The death of the dark room is not a claim that all AI has become transparent. It is a claim that boards' minimum acceptable standard of explainability has been durably raised.

This argument holds most strongly for board-salient, horizontally comparable outputs such as chat, summarisation, search, drafting, and content generation. That is where the consumer benchmark bites hardest. More complex deployments – agentic systems, specialised vertical models, workflow orchestration, predictive systems buried in process – can rebuild opacity at another layer. HiddenLayer, SAS, and McKinsey have all pointed to the explainability problems emerging in agentic architectures.

So the door does not reopen in its old place. Instead, new dark rooms form elsewhere.

One is the vendor dark room, where weights, training data, evaluation methods, and inference logic sit behind commercial confidentiality. Another is the governance dark room, where oversight committees, policies, and dashboards reintroduce asymmetry one level above the technical function. A third is the shadow AI dark room. Harmonic Security's 2026 research, based on more than 22 million enterprise prompts, found extensive use of personal AI tools and material exposure of sensitive data through personal accounts. VentureBeat, drawing on Ivanti data, described the same migration from sanctioned systems to personal ones when enterprise tools create too much friction.

When visibility becomes excessive, workers route around it.

This is where Bernstein's transparency paradox matters. His setting was manufacturing rather than enterprise AI, so the transfer is analogical rather than direct. But the mechanism travels well: extreme observability changes behaviour, and not always for the better. People start performing for the metric, the meeting, or the demo.

That is why the death of the dark room creates a new capital allocation bias. Demo-friendly, screenshot-ready use cases rise quickly because they are politically legible. Slower back-office transformations, the ones that compound over years, lose oxygen because they do not photograph well. Political visibility and economic value are not the same thing.

## The new mandate

*A board member who cannot read an architecture diagram can forward a screenshot of a bad AI response.*

That single change moves the CIO role from Chief Engineer of complexity to Chief Actuary of visible risk. The job is no longer only to make systems function. It is to manage what becomes politically visible, to whom, in what sequence, with what protection, and with what explanation ready before the first screenshot travels upward.

Whether the function is called CIO, Chief AI Officer, or board committee does not change the mandate. Someone must manage the political consequences of AI visibility. For horizontal AI, boards can judge outputs directly. For more complex systems, they still rely on technical expertise. But even there the burden of proof has shifted, because boards that can judge some AI outputs now expect justification for all of them.

That is also where the next questions in this series begin. Once a deployment becomes visible, the liability question hardens. Once the board can see the risk, it starts asking who owns it. Those are separate briefings. This one is about the moment before them, when the room goes bright.

## Three questions to ask before the formal governance reporting starts flattering you

1. Can your board independently verify the performance of your most visible AI deployment against a consumer alternative?
2. If your most successful AI pilot became fully visible tomorrow, which roles, committees, or empires would it threaten?
3. Are you selecting AI investments for economic value or for political legibility – and inside your organisation, who can still tell the difference?

## Monday morning

1. Map which AI deployments are already visible to non-technical stakeholders and which ones are about to become visible.
2. Identify which organisational roles are most threatened by visible AI success, not just by visible AI failure.
3. Audit whether your governance model still assumes opacity for outputs the board can now judge directly.
4. Separate demo-friendly AI from economically consequential AI in your investment portfolio before board attention does it for you.

The death of the dark room marks the end of a thirty-year governance equilibrium. When AI effectiveness is instantly visible and every deployment is politically consequential, the CIO stops being the company's Chief Engineer of complexity. They become its Chief Actuary of visible risk – responsible not for what the technology computes, but for what the organisation can no longer unsee.

## References

- Ben-Zur, L. (2025, June 29). The innovation paradox: Why organizations resist the future they say they want. *LBZ Advisory*. <https://liatbenzur.com/2025/06/29/innovation-paradox-why-organizations-resist-future-they-want/>
- Bernstein, E. S. (2012). The transparency paradox: A role for privacy in organizational learning and operational control. *Administrative Science Quarterly*, 57(2), 181–216. <https://doi.org/10.1177/0001839212453028>
- Birkinshaw, J., & Ridderstråle, J. (1999). Fighting the corporate immune system: A process study of subsidiary initiatives in multinational corporations. *International Business Review*, 8(2), 149–180. [https://doi.org/10.1016/S0969-5931\(98\)00043-2](https://doi.org/10.1016/S0969-5931(98)00043-2)
- Boston Consulting Group. (2025, September). *The widening AI value gap: Build for the future 2025*. <https://www.bcg.com/publications/2025/are-you-generating-value-from-ai-the-widening-gap>
- Brown, I. (2025, March 24). Beyond the black box: How agentic AI is redefining explainability. *SAS Voices*. <https://blogs.sas.com/content/sascom/2025/03/24/beyond-the-black-box-how-agentic-ai-is-redefining-explainability/>
- Challapally, A., Pease, C., Raskar, R., & Chari, P. (2025, July). *The GenAI divide: State of AI in business 2025*. MIT Media Lab, Project NANDA.
- Columbus, L. (2025, September 20). Legacy UI is dead: Shadow AI is how real work gets done now. *VentureBeat*. <https://venturebeat.com/security/legacy-ui-is-dead-shadow-ai-is-how-real-work-gets-done-now/>
- Dataiku & Harris Poll. (2026). *The 7 career-making AI decisions for CIOs in 2026* [Survey report]. <https://pages.dataiku.com/cio-ai-decisions>
- Deloitte China & HKU Centre for AI, Management and Organisation. (2026, January 16). *AI Adoption Index 2026: The paradox of promise and performance* [Research report]. <https://camo.hku.hk/ai-adoption-survey/>

- Dyki, K. (2026). *Selling surplus or scarcity: Where AI margin actually lives* [Strategic briefing]. Quasen. <https://doi.org/10.5281/zenodo.18789037>
- Eatough, E., Ferrazzi, K., Smith, W., & Waters, S. (2026, February 17). Why AI adoption stalls, according to industry data. *Harvard Business Review*. <https://hbr.org/2026/02/why-ai-adoption-stalls-according-to-industry-data>
- Falconer, S. (2026, January 27). Stop treating enterprise AI like consumer AI. *Medium*. <https://seanfalconer.medium.com/stop-treating-enterprise-ai-like-consumer-ai-7acd256bce78>
- Harmonic Security. (2026, January 15). *What 22 million enterprise AI prompts reveal about shadow AI in 2025* [Research report]. <https://www.harmonic.security/resources/what-22-million-enterprise-ai-prompts-reveal-about-shadow-ai-in-2025>
- Hermann, E., Puntoni, S., & Morewedge, C. K. (2026, March–April). Why gen AI feels so threatening to workers. *Harvard Business Review*. <https://hbr.org/2026/03/why-gen-ai-feels-so-threatening-to-workers>
- Hilb, M. (2025, August). From information asymmetry to intelligence symmetry: How AI will reshape corporate governance. *Board Foundation*. <https://boardfoundation.org/en/insight/from-information-asymmetry-to-intelligence-symmetry-how-ai-will-reshape-corporate-governance/>
- ISS-Corporate. (2025, March 19). Roughly one-third of large U.S. companies now disclose board oversight of AI. *ISS Insights*. <https://insights.issgovernance.com/posts/roughly-one-third-of-large-u-s-companies-now-disclose-board-oversight-of-ai-iss-corporate-finds/>
- Jensen, M. C., & Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics*, 3(4), 305–360. [https://doi.org/10.1016/0304-405X\(76\)90026-X](https://doi.org/10.1016/0304-405X(76)90026-X)
- McKinsey & Company. (2025, December 4). The AI reckoning: How boards can evolve. <https://www.mckinsey.com/capabilities/mckinsey-technology/our-insights/the-ai-reckoning-how-boards-can-evolve>
- McKinsey & Company. (2025, October 20). The new economics of enterprise technology in an AI world. <https://www.mckinsey.com/capabilities/mckinsey-technology/our-insights/the-new-economics-of-enterprise-technology-in-an-ai-world>
- Monnette, J., & Chaudhary, A. (2025, October 27). Why do 80% of AI pilots fail to scale? Unpacking the top enterprise AI deployment challenges. *EPAM Systems*. <https://www.epam.com/insights/ai/blogs/enterprise-ai-deployment-challenges>
- Nolan, R. L., & McFarlan, F. W. (2005). Information technology and the board of directors. *Harvard Business Review*, 83(10), 96–106.
- Pearcy, S. (2025, July 9). Structuring transparency for agentic AI. *HiddenLayer*. <https://www.hiddenlayer.com/insight/structuring-transparency-for-agentic-ai>
- Pfeffer, J., & Salancik, G. R. (1978). *The external control of organizations: A resource dependence perspective*. Harper & Row.
- S&P Global Market Intelligence / 451 Research. (2025, October). Generative AI shows rapid growth but yields mixed results. <https://www.spglobal.com/market-intelligence/en/news-insights/research/2025/10/generative-ai-shows-rapid-growth-but-yields-mixed-results>
- Sharma, A. (1997). Professional as agent: Knowledge asymmetry in agency exchange. *Academy of Management Review*, 22(3), 758–798.

## Cite This Paper

### APA 7:

Dyki, K. (2026). *Death of the dark room: How generative AI broke enterprise IT's political cover* (Strategic Briefing). Quasen. <https://doi.org/10.5281/zenodo.18961987>.

### **BibTeX:**

```
@techreport{dyki2026darkroom,  
  author      = {Dyki, Krzysztof},  
  title       = {Death of the Dark Room: How Generative AI Broke Enterprise IT's Political  
Cover},  
  institution = {Quasen},  
  year        = {2026},  
  month       = {March},  
  type        = {Strategic Briefing},  
  doi         = {10.5281/zenodo.18961987},  
  url         = {https://doi.org/10.5281/zenodo.18961987}  
}
```

### **Chicago (Author-Date):**

Dyki, Krzysztof. 2026. "Death of the Dark Room: How Generative AI Broke Enterprise IT's Political Cover." Strategic Briefing, Quasen. <https://doi.org/10.5281/zenodo.18961987>.