InfoSecAI

A BRIEF FOR CISOS · BISOS · GRC LEADERSHIP

WHITE PAPER · PAPER 2 · 2026

The Shadow AI Exposure Map

A practitioner's brief for CISOs, CIOs, IT risk leaders, data protection officers and procurement leaders mapping the AI estate they did not know they had.

AUTHORED BY

Paul Jolliffe

By Paul Jolliffe, Founder and Director, InfoSecAI Limited · MBA · CISSP · ISO 27001 Lead Auditor

01 · EXECUTIVE SUMMARY

Executive summary

In most organisations the artificial intelligence (AI) estate is larger than the inventory. Browser-based chatbots, embedded AI features in everyday software-as-a-service (SaaS) tools, copilots wired into productivity suites, and departmental pilots procured outside the central technology budget are all in use. Few have been registered, classified, controlled or evidenced.

That is shadow AI. It is not primarily a behaviour problem. It is a visibility and enablement problem. Employees are not being reckless; they are being productive. The governance gap is structural, not cultural.

This brief is the second in InfoSecAI's five-part executive series, From AI Ambition to AI Assurance. It sets out where shadow AI hides, why acceptable-use policies do not see it, and how to build a Shadow AI Exposure Map that turns an unknown estate into a controlled one in thirty days. The discipline is secure enablement, not blanket restriction. Banning AI tends to push it deeper into the shadow; meeting demand with sanctioned options pulls it back into the light.

The first paper in this series argued that AI governance is no longer a policy problem; it is an operating-model problem. This paper extends that argument into the discovery activity at the front of the operating model. The AI system register cannot be operated against an estate the organisation cannot see.

02 · WHY THIS MATTERS NOW

Why this matters now

Three forces are widening the gap between sanctioned AI and actual AI use.

The first is supplier-embedded AI. Generative AI (GenAI) features are now bundled into most enterprise software-as-a-service (SaaS) contracts. Customer relationship management platforms, productivity suites, project management tools, design software, code editors, marketing automation and helpdesk systems all ship AI capabilities by default. Many were enabled in mid-2025 platform updates without explicit organisational opt-in. The supplier disclosed the change in release notes; the procurement file did not.

The second is employee-led adoption. Public chatbots, browser extensions and free-tier AI tools are zero-friction to start and demonstrably faster at common knowledge work. Employee use runs ahead of formal organisational deployment in most functions, most acutely in legal, marketing, customer success and software engineering. The pattern tracks the earlier consumerisation of IT wave, with one difference: AI tools can ingest sensitive data in volume in a single paste.

The third is the regulatory clock. The European Union Artificial Intelligence Act (EU AI Act, Regulation (EU) 2024/1689) applies the bulk of its high-risk system obligations from 2 August 2026, including the deployer obligations under Article 26 for third-party AI. The

Information Commissioner's Office (ICO) in the United Kingdom signals the same accountability principle. "We did not know our supplier added that AI feature" does not survive a supervisory request.

03 · WHY CURRENT APPROACHES ARE FAILING

Why current approaches are failing

Three control patterns recur, each individually sensible, all three together insufficient.

The first is the acceptable-use policy. Necessary, but not a control. A policy detects nothing, logs nothing and triggers no alert when ignored. Without a discovery layer, there is no evidence of compliance and no evidence of breach.

The second is the platform block. Blocking specific public AI tools at the proxy catches the most visible chatbots. It does not catch AI features embedded in already-allowed SaaS, AI inside browser extensions, AI on personal devices or AI accessed through the same vendor's separately approved service. Blocking is a partial control with a high false sense of completeness.

The third is the procurement gate. A new vendor contract triggers a review; an AI feature added mid-contract often does not. The largest shadow AI exposures sit inside platforms already trusted with the relevant data.

The structural fix is not a longer policy. It is to add discovery, classification and enablement as continuous activities, not one-off projects.

04 · WHERE SHADOW AI HIDES

Where shadow AI hides

Five hiding places dominate in the organisations InfoSecAI reviews. The list is not exhaustive; it is the starting search space for a discovery exercise.

The first is supplier-embedded AI. Generative AI features inside enterprise SaaS, often introduced via release-note updates rather than contract amendments. These are the lowest-friction exposures and the highest-volume ones, because the SaaS tool already holds organisational data.

The second is browser-based AI. Public chatbots accessed from corporate or personal browsers, browser extensions that summarise pages or rewrite text using third-party models, and AI-enhanced search experiences. The data flows are easy to underestimate because they look like ordinary web traffic.

The third is developer-tooling AI. Code assistants, AI-augmented integrated development environments, AI commit summarisers, AI security review tools. Each can ingest proprietary source code, secrets or customer data depending on how the integration is configured.

The fourth is departmental pilots. Marketing, sales, legal, customer success and human resources teams running AI tools procured on a corporate card or via a free tier. These are often the most directly value-creating uses and the least visible to central technology and security functions.

The fifth is agentic AI. Tools that go beyond content generation into action: scheduling, sending messages, retrieving data, taking decisions or triggering workflows. Agentic AI is the subject of the next paper in this series. Even pre-deployment, organisations should treat any agent in evaluation as part of the shadow AI estate until classified.

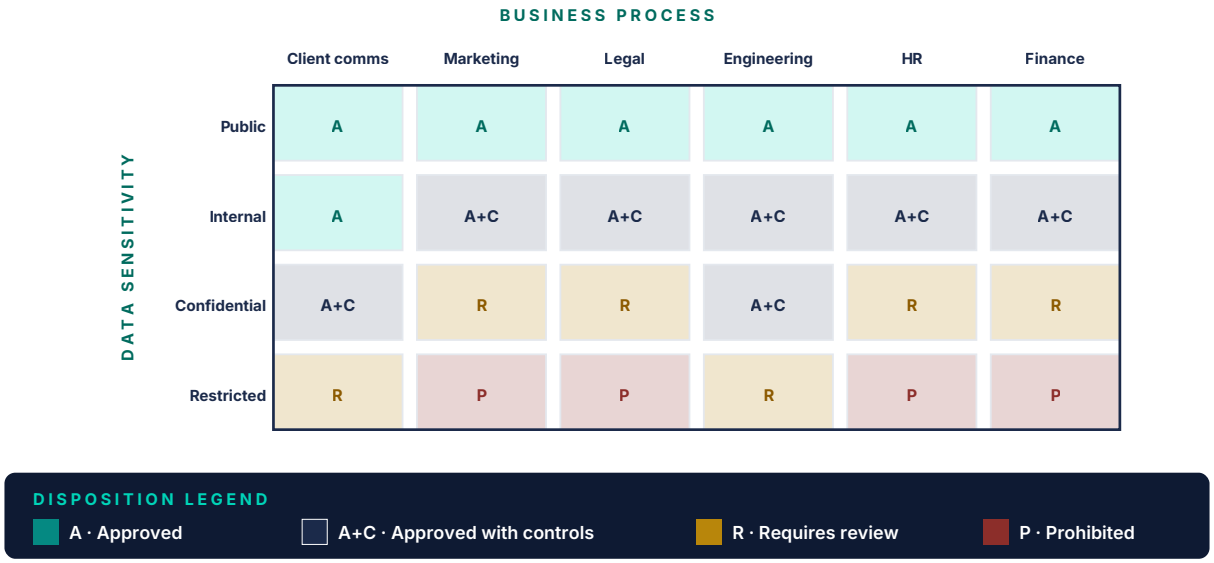
05 · THE SHADOW AI EXPOSURE MAP

The Shadow AI Exposure Map

The map below is the discovery and classification artefact that turns an unknown estate into a controlled one. It is built around four facts per AI use case: the business process it touches, the data sensitivity it consumes, the supplier or model behind it, and the current approval status.

FIGURE 1 · The Shadow AI Exposure Map.

Business process × data sensitivity. Cells coloured by disposition. Refreshed monthly.



The map is operated, not produced. The first version is built in a thirty-day discovery sprint described below. The continuing operation is monthly: new SaaS features added by suppliers, new tools requested by business teams, and new use cases approved through the intake process all flow into the same register.

Each row carries one decision: what disposition is the organisation taking against this use case. There are four dispositions: approved, approved with controls, requires review, prohibited. The disposition is set by the named control owner using the use-case classification matrix introduced in Paper 1 of this series and the supplier review checklist referenced below.

06 · THE AI DEMAND SIGNAL

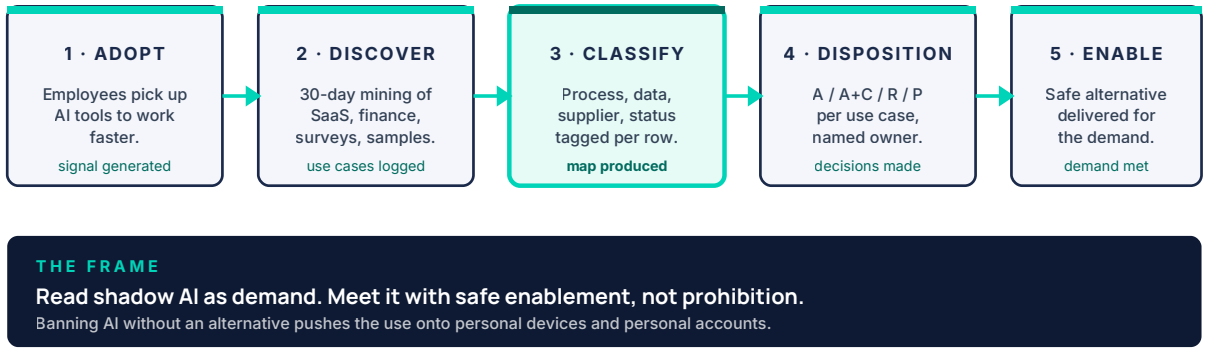
The AI demand signal

The framing matters. Shadow AI is best read as an AI demand signal, not a compliance failure.

When 60 per cent of a sales team is using a public chatbot to draft proposals, the signal is not "the sales team is bad at policy". The signal is "there is unmet demand for AI-assisted proposal drafting and the sanctioned options do not exist or are not usable". The governance response of value is to look at what employees are using the tool for, classify the use cases, and provide a sanctioned alternative for the use cases that are safe to enable.

FIGURE 2 · The AI demand signal flow.

Shadow AI is unmet demand surfacing. Read it as a signal, not a failure.



The discipline is to treat each shadow AI use case as both a risk to assess and a product requirement to meet. Three of the five typical hiding places listed earlier are amenable to enablement: supplier-embedded features can be configured for data minimisation, browser-based AI can be replaced with an enterprise-licensed alternative on the corporate proxy, and departmental pilots can be brought into the intake process with the named owner unchanged.

Banning AI without offering a sanctioned alternative tends to push the use deeper into the shadow. Employees move to personal devices, personal accounts and tools the corporate proxy cannot see. The risk profile gets worse, not better.

07 · THE FOUR DISPOSITIONS

The four dispositions

Every discovered AI use case is assigned one of four dispositions. The decision is made within five working days of discovery, by the named control owner, using documented criteria.

Approved. The data sensitivity is low, the supplier holds the appropriate processing terms, and the AI feature is configured for the organisation's data minimisation expectations. Examples: public-domain document summarisation, marketing image generation against non-confidential briefs.

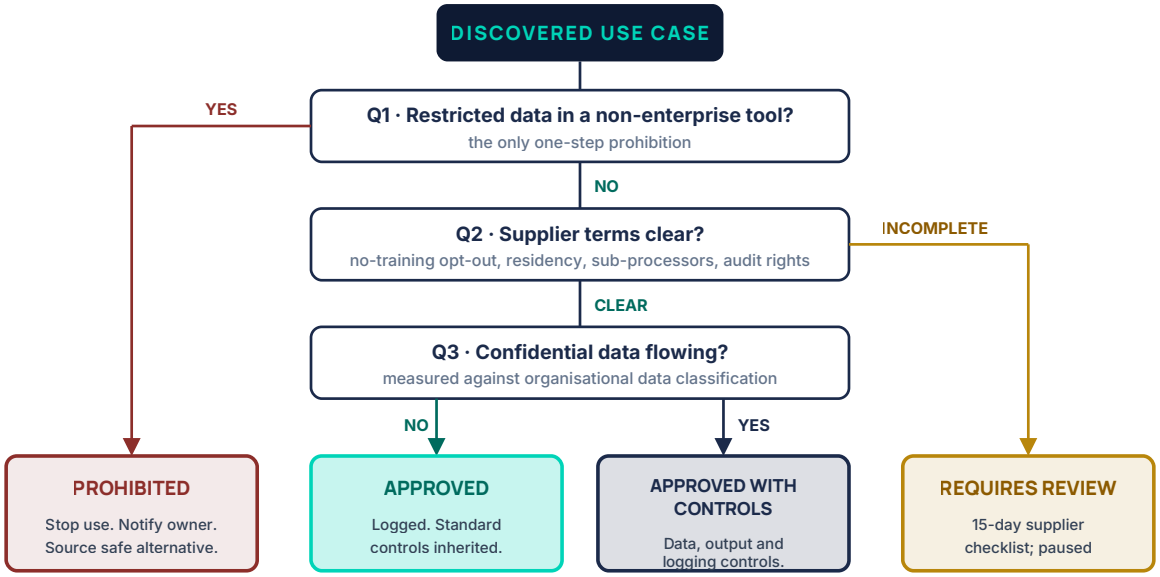
Approved with controls. The use case is acceptable subject to specified controls: data classification training, output review before external use, prompt filtering, logging to a centralised store. Most enterprise SaaS-embedded AI features land here.

Requires review. The use case is potentially valuable but the risk profile is unclear or the supplier disclosures are incomplete. The control owner triggers a structured review using the SaaS AI supplier checklist. The use case is paused, not killed, while the review runs. Maximum review duration is fifteen working days; longer triggers an escalation.

Prohibited. The use case cannot be made safe under the organisation's current control framework. Examples: pasting restricted personal data into public-tier consumer chatbots, AI tools whose suppliers reserve the right to train on customer inputs without an enterprise opt-out, AI used to make individual decisions about employees or customers without the human oversight expected under EU AI Act Article 14.

FIGURE 3 · The four-disposition decision tree.

Discovered AI use case in. One of four dispositions out. Owner-overridable with documented justification.



Override authority sits with the named control owner with documented justification. The tree is the default; it is not a substitute for judgement on edge cases.

08 · THE FIRST THIRTY DAYS OF DISCOVERY

The first thirty days of discovery

A working Shadow AI Exposure Map can be produced in thirty calendar days without procuring a discovery tool. Tooling helps later. The first version uses sources the organisation already has.

In week one, mine the procurement, finance and SaaS administration data. Pull supplier contracts and release notes for the top fifty SaaS tools, flag every product that introduced AI features in the last twenty-four months, and add each as a row. Cross-reference against the existing IT asset register and the data-classification scope. Target completeness: 90 per cent of supplier-embedded AI exposures.

In week two, run a structured discovery survey across the business. Ask each function: which AI tools are in use, for what purpose, against what data, paid or free. Frame the survey as productivity research, not compliance audit. The response rate matters more than the wording; the question being asked openly is the signal.

In week three, sample three high-traffic functions for browser and developer tooling AI. Marketing, sales and engineering produce the most data per minute and are the most likely to have adopted browser-based AI. A two-day desk review of recent outputs and a short interview with each team lead surfaces the bulk of the in-use tools.

In week four, classify and disposition. Run every discovered row through the four-disposition tree. Produce the v1 map. Walk the map with the executive sponsor and the three highest-exposure function owners. The output is a classified register and a list of the five use cases needing safe-enablement alternatives.

By the end of week four the organisation has an exposure map, four dispositions assigned, a short list of enablement priorities, and a documented control gap for the cases that need it. That is not a finished programme. It is a foundation on which the AI control evidence register from Paper 1 can be operated.

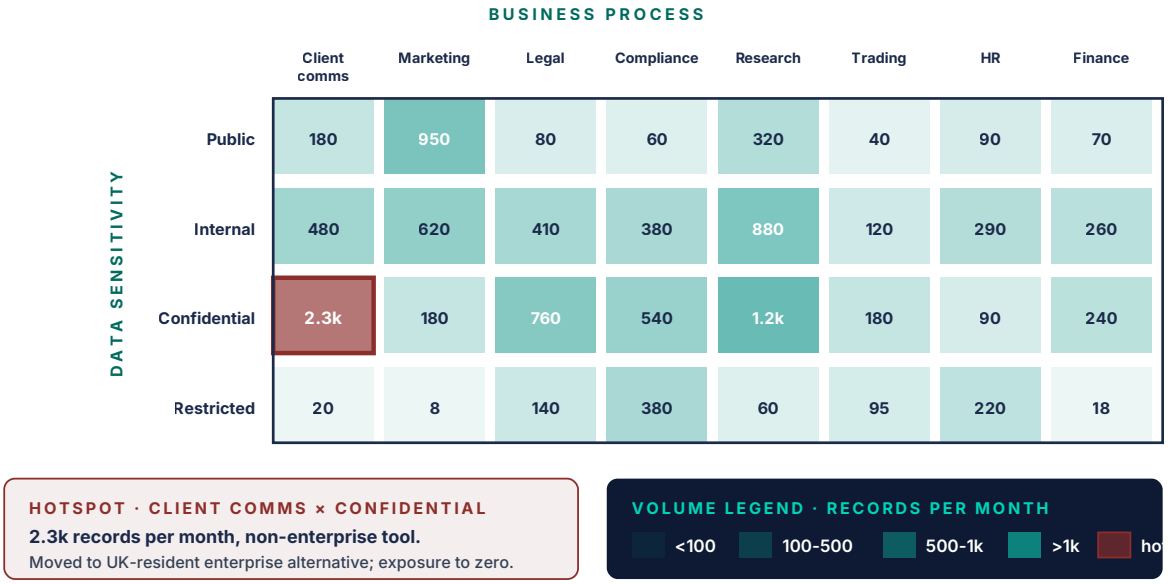
09 · THE DATA EXPOSURE HEATMAP

The data exposure heatmap

A useful side artefact of the discovery sprint is a data exposure heatmap: business processes by data sensitivity, with cells coloured by the volume of AI processing flowing through them. The heatmap turns the register into a one-page executive view.

FIGURE 4 · Data exposure heatmap.

Worked example: Pennycross Capital Partners plc, April 2026. Cell value = AI-processed records per month.



Worked example. Pennycross Capital Partners plc, a mid-sized UK asset manager, ran the discovery sprint in April 2026 with a CISO, a procurement lead and a data protection officer (DPO). The discovery surfaced 38 distinct AI use cases across the firm. Of those, 14 were sanctioned and well-controlled; 16 were sanctioned in principle but lacked documented controls; 6 were genuinely shadow, including a marketing-team browser extension summarising client briefings into a third-party model with non-enterprise terms; 2 were prohibited and stopped within ten working days. The resulting safe-enablement programme moved the client-briefing summarisation to an enterprise-licensed alternative with data residency in the United Kingdom and an explicit no-training clause. Reported productivity gain in the marketing team was unchanged; data exposure dropped to zero for the controlled flow.

The map is the entry point. The enablement is the conversion. The combination is what turns shadow AI from a compliance worry into a productivity programme with controls.

10 · INFORMATION SECURITY IMPLICATIONS

Information security implications

Six integration points matter, extending the security model from Paper 1 into the discovery surface.

Data classification training is the prerequisite control for any "approved with controls" disposition. If employees cannot reliably tell restricted from confidential from internal, paste-based AI use leaks across the boundary regardless of policy.

Network and proxy visibility matters less than expected. Most shadow AI sits inside permitted SaaS or runs through permitted domains. Proxy logs are useful evidence after the fact, not the discovery layer.

Supplier contracts carry the largest exposure category. The minimum clauses are: explicit opt-out from training, data residency, sub-processor disclosure, breach notification timelines, audit rights for AI processing specifically, and an exit obligation that includes deletion of derived embeddings.

Output handling controls reduce the risk of AI content being treated as authoritative. The Open Worldwide Application Security Project (OWASP) Top 10 for Large Language Model Applications (2025 edition) lists sensitive information disclosure, output handling and prompt injection in its top five. Practical controls: review before external publication, source attribution for AI-assisted research, and a documented prohibition on using AI outputs as the sole basis for decisions about individuals.

Logging at prompt and output level for sanctioned tools is the evidence base for assurance. EU AI Act Article 12 mandates retention for high-risk systems; six months is the pragmatic minimum for shadow-AI-derived prompts.

Incident response needs an explicit shadow-AI branch in the classification standard operating procedure (SOP). The trigger is "AI processed data it should not have, or produced output that was acted upon when it should not have been". Both paths need playbooks.

11 · EXECUTIVE DECISION POINTS AND QUESTIONS

Executive decision points and questions

Six executive decisions reliably unblock the largest exposures.

Who is accountable for the AI estate inventory? One named individual, typically the AI governance lead, with the chief information security officer (CISO) as accountable executive.

What is the organisation's position on supplier training-data use? The default should be no training without explicit opt-in by the named control owner, regardless of contract age.

What is the default disposition for confidential or restricted data flowing through AI? Approved with controls; prohibited for restricted data flowing through non-enterprise tooling.

What is the safe-enablement budget? The cost of moving the top five shadow uses to sanctioned alternatives. Typically modest; almost always less than the first material incident.

What is the refresh cadence? Monthly for high-velocity functions; quarterly for the wider organisation; ad-hoc on major SaaS renewals.

What evidence proves the inventory is current? A map dated within thirty days, the change log against the prior version, and a named owner per row.

The questions every leader should be ready to answer this quarter. Can you produce the current AI inventory in under one hour? How many use cases moved in the last thirty days? For the three highest-volume flows, do you know the supplier's training-data

position, the data residency and the named owner? What is the safe-enablement gap, expressed as the number of shadow uses awaiting a sanctioned alternative? That number is the priority list.

12 · CLOSING THOUGHT

Closing thought

Shadow AI is not a behaviour problem. It is a visibility, control and enablement problem. The discovery activity is the front of the AI governance operating model; without it, every downstream control is operating against an estate the organisation cannot see.

The thread of this series continues: AI assurance evidence, not AI reassurance narrative. An organisation that knows its AI estate can defend it. An organisation that does not, cannot. The discipline is to meet employee demand with safe enablement rather than meet it with prohibition.

The next paper, Securing Agentic AI Before It Acts, examines what changes when AI moves from generating content to taking action.

13 · SOURCE REGISTER

Source register

All sources verified to primary publisher on 2 June 2026.

#	SOURCE	USE IN PAPER	LINK
1	Regulation (EU) 2024/1689 (the EU AI Act)	Article 26 deployer obligations, Article 14 human oversight, Article 12 logging, applicability timeline	https://eur-lex.europa.eu/eli/reg/2024/1689
2	NCSC, Guidelines for secure AI system development	Secure deployment and operation framing for the discovery and control layers	https://www.ncsc.gov.uk/collection/guidelines-secure-ai-system-development
3	OWASP Top 10 for Large Language Model Applications, 2025	Prompt injection, sensitive information disclosure, supply chain, output handling	https://genai.owasp.org/llm-top-10/
4	ISO/IEC 42001:2023, AI Management System Standard	Inventory and classification activities in Annex A controls	https://www.iso.org/standard/42001
5	NIST AI Risk Management Framework (AI RMF 1.0)	Govern and Map function framing for the discovery activity	https://www.nist.gov/itl/ai-risk-management-framework

#	SOURCE	USE IN PAPER	LINK
6	ICO AI guidance and AI audit framework	UK supervisory expectation that controllers know the AI estate	https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/artificial-intelligence/
7	European Commission, AI Act regulatory framework	Provider versus deployer role determination	https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai

14 · ABOUT THIS SERIES

About this series

From AI Ambition to AI Assurance is a five-paper executive briefing series, 1 to 5 June 2026.

1. AI Governance Is No Longer a Policy Problem
2. The Shadow AI Exposure Map (this paper)
3. Securing Agentic AI Before It Acts
4. Why AI Transformation Fails After the Pilot
5. The Board Pack for AI Assurance

Each paper is published as a 12-page executive briefing under the InfoSecAI Blog Template. The full series is available at infosecai.net/insights for subscribers to the InfoSecAI insights list.

15 · PRACTITIONER NOTE

Practitioner note

This briefing is practitioner interpretation, not legal advice. For regulated deployments, validate final claims against current legal obligations, sector-specific requirements and the original primary sources before relying on them.

InfoSecAI

About InfoSecAI

InfoSecAI is an independent UK consultancy helping organisations turn security, regulatory, resilience and AI governance requirements into practical operating models, stronger controls and robust delivery.

We work across strategy, governance, risk, compliance, AI security, assurance, operations and engineering. Our services help leadership teams assess their current position, align to standards and regulation, define the target operating model, and deliver the governance, controls, artefacts and ways of working needed to move from intent to implementation.

Our toolkit capability accelerates structured work across ISO 27001, ISO 22301, ISO 42001, NIST CSF, NIST AI RMF, CIS Controls, Cyber Essentials, DORA, NIS 2, the EU AI Act, GDPR, UK GDPR, SOC 1 and SOC 2. The approach combines AI-enabled workflow support with senior practitioner judgement, so outputs remain proportionate, usable and connected to the way the organisation actually operates.

InfoSecAI was founded in **2025** by **Paul Jolliffe**. The company is built for organisations that need clarity, senior leadership and hands-on delivery across information security and AI governance, without adding unnecessary complexity or treating compliance as a paperwork exercise.

infosecai.net · paul.jolliffe@infosecai.net