AI Discovery

Agent Tool Interoperability: Standards Boards Should Care About

When every suite ships an agent, the scarce asset becomes controlled tool access to your systems of record, and the ability to switch orchestration without rewiring the enterprise. Boards should care about interoperability as procurement hygiene, not as developer tribalism. The issue is not which acronym wins. It is whether identity, audit, and exit options remain under your control.

David Saliba . 2026-07-17 . aimonger.com/whitepapers/enterprise-agent-tool-interoperability/

The problem in one sentence

Every major software suite will ship an agent that wants credentials to your ERP, CRM, document store, and ticketing system. If each agent gets its own integration path, you lose auditability, multiply privilege surfaces, and pay twice to exit a vendor.

Agent, plainly: software that chooses actions (often via tools) toward a goal under policy limits, not a chat window that only drafts text. **In practice:** an ITSM agent that opens, updates, and closes tickets through approved APIs, with every call logged, not a Copilot pane that pastes ticket text for a human to retype.

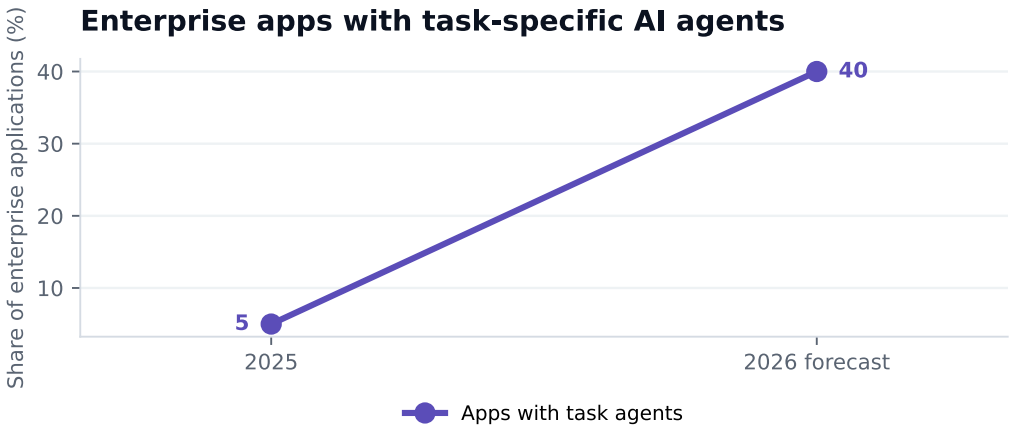
Tool interoperability, in short: the ability to invoke the same governed business function (create ticket, fetch contract, post journal) from more than one orchestrator without duplicating credentials or rewriting schemas. **Worked case:** when the CRM contract renews at 40% uplift, you can move ticket automation to an internal gateway without a six-month re-integration project.

From model choice to tool choice

The model layer commoditises. The tool layer, covering ERP, CRM, DMS, ticketing, and identity, is where the enterprise actually lives. Agents that cannot safely use tools are toys. Agents that use tools without interoperable control planes become silent lock-in.

In a mid-market firm, this becomes visible quickly. The productivity suite wants credentials, the CRM wants credentials, a specialist agent wants credentials, and the security team is asked to approve three different ways of touching the same ticketing or document system.

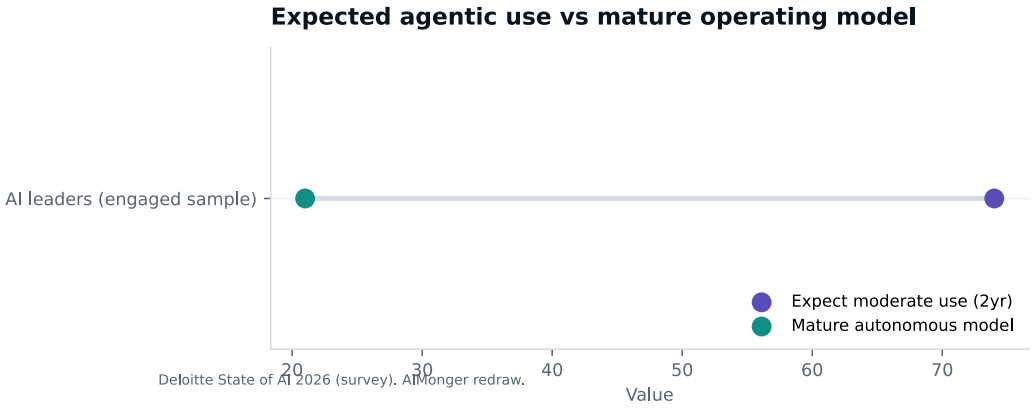
What you see globally right now is agent proliferation inside software suites. Gartner predicts that by the end of 2026, 40 per cent of enterprise applications will feature task-specific AI agents, up from less than 5 per cent in 2025. When agent access spreads faster than tool control, lock-in and privilege sprawl arrive before the business case is proven.



Gartner press release, 26 Aug 2025 (forecast). AIMonger redraw.

Figure 1. Share of enterprise applications expected to include task-specific AI agents. Plain read: agents become a default product checkbox, not a rare pilot. **Operating case:** by 2026 your CRM, ERP connector, and productivity suite may each ship an agent that requests ITSM credentials separately. Source: Gartner, "40% of Enterprise Apps Will Feature Task-Specific AI Agents by 2026," press release, 26 August 2025 (https://www.gartner.com/en/newsroom/press-releases/2025-08-26-gartner-predicts-40-percent-of-enterprise-apps-will-feature-task-specific-ai-agents-by-2026-up-from-less-than-5-percent-in-2025). AIMonger redraw.

Deloitte's 2026 enterprise AI survey reports that 74 per cent of AI leaders expect at least moderate agentic use within two years, while only 21 per cent describe a mature operating model for autonomous agents. That gap is an interoperability gap: many orchestrators, few governed tool boundaries.



Deloitte State of AI 2026 (survey). AIMonger redraw.

21 74

Figure 2. AI leaders expecting moderate agentic use within two years vs those with a mature autonomous-agent operating model. In short: suites will ship agents before your estate is ready to govern them. **On Monday:** the board sees three vendor demos; only one has a tool registry and audit export. Source: Deloitte State of AI in the Enterprise 2026, survey of 3,235 leaders (survey). AIMonger redraw.

Orchestrator means: the runtime that plans steps, calls models, and invokes tools, whether vendor-hosted or internal. **Operating case:** Microsoft Copilot Studio, Salesforce Agentforce, or an internal LangGraph service are all orchestrators; the question is whether they share tool boundaries.

Plain read of the evidence: access to agents is becoming mainstream; governed tool access is not. Procurement must shift from "which agent brand" to "which tool contract."

Board procurement questions

Before signing or renewing suite agent SKUs, the CIO should be able to answer yes to each:

1. **Portable tool invocation.** Can we invoke the same tools from an alternative orchestrator through our gateway or documented API? This question tests whether the vendor has welded its agent to proprietary connectors, or whether the underlying business functions remain reachable from a second runtime. A yes means exit cost stays bounded; a no means the workflow is hostage to the renewal calendar.

2. **Enterprise identity and immutable logging.** Is every tool call authenticated with enterprise identity and logged immutably? Without identity federation, agents run on shared service accounts that nobody can revoke selectively, and without immutable logs, incident review depends on vendor screenshots. Both gaps fail EU deployer duties when the answer influences a consequential decision.

3. **Trace export for evaluation and FinOps.** Can we export traces for evaluation, FinOps attribution, and incident review? Replayable JSON traces are the only artefact that lets a red-team reproduce a failure and lets finance attribute token spend to a workflow. Vendor dashboards that show aggregate counts without raw export are evidence of nothing.

4. **Data boundary and residency.** What data leaves our boundary on each tool call, and under which residency terms? Agents that forward prompt context or tool payloads to a vendor region outside the contract create a data-processing event the firm may not have authorised. The answer must name the region and the subprocessors, not "best efforts."

5. **Exit plan on deprecation or rebrand.** What is the exit plan if the agent suite is deprecated, repriced, or agentwashed into a rebrand? Gartner forecasts heavy agentic cancellation by 2027, so some SKUs will end. The contract should commit to connector documentation, bulk trace export, and a transition window, not silent end-of-life.

6. **Agent versus assistant labelling.** Does the vendor distinguish assistants (draft-only) from agents (tool use under policy)? The distinction drives which controls apply: a draft assistant needs logging, a tool-using agent needs allowlists, budgets, and gates. SKUs that blur the line let vendors sell assistants at agent prices while buyers apply agent controls to harmless drafting tools.

Agentwashing, put simply: marketing that labels a scripted assistant or RPA wrapper as an "agent" without material planning, tool policy, or autonomy controls. **Board reading:** a vendor demo shows "autonomous IT resolution" but cannot export audit logs or define stop conditions; that is agentwashing, and Gartner has publicly warned the market about it.

Evidence base: proliferation precedes control

Source	Finding	What it proves	What it does not prove
Gartner (2025)	40% of enterprise apps expected to feature task-specific agents by end-2026	Multiple suite agents will request tool credentials	That any single vendor stack will dominate
Gartner (2025)	>40% of agentic AI projects forecast cancelled by end-2027	Weak cost/risk controls kill programmes	That cancellation is random: lock-in and missing metrics are named drivers
Deloitte 2026	74% expect moderate agent use; 21% mature operating model	Interoperability hygiene must precede sprawl	That Deloitte respondents represent all EU mid-market firms
OWASP Agentic Top 10 2026	Tool misuse, privilege abuse, supply chain, insecure inter-agent comms	Tool registries and scoped credentials are mandatory	That any protocol alone removes these risks
FinOps 2026	Nearly all surveyed practitioners manage AI spend	Tool calls are cost events needing stable IDs	Exact attribution methods: buyer must define

McKinsey's State of AI 2025 survey finds that high performers, organisations attributing more than 5 per cent of EBIT to AI, scale what works with senior ownership and workflow redesign. Interoperability is part of that scaling story. If every workflow is welded to one vendor orchestrator, scaling stops at the contract boundary.

Practical posture: own the spine, rent the orchestrator

Technology leaders should assume a multi-vendor agent estate through 2028:

- **Own identity and authorisation centrally.** Route every agent through your identity provider using OIDC or OAuth patterns, never shared service accounts scattered across agents. Central ownership means one revoke point in an incident, and it stops vendors from accumulating duplicate credentials to the same system of record. A Milan insurer we worked with in principle issues short-lived tokens per agent case, so a compromised CRM agent cannot touch the claims system an hour after the breach is found.
- **Prefer tools wrapped behind an internal gateway.** Where blast radius or duplication cost justifies platform engineering, place a governed front door between agents and systems of record so schemas are enforced and budgets apply. The gateway becomes the single place where allowlists, rate limits, and audit events live, which is cheaper than reconciling three vendor audit formats after an incident.
- **Require evaluation harness access outside the vendor UI.** Contracts must guarantee trace export, replay, and red-team cases that run independently of the vendor console. A vendor dashboard that shows green is not evidence; replayable JSON traces that feed your own eval pack are. Without this clause, the first prompt-injection failure will be unattributable.
- **Keep workflow metrics independent of orchestrator brand.** Track cost per successful case, tool pass rate, and incident count as workflow properties, not as properties of whichever agent platform happened to run them. These metrics let you compare a suite agent against an internal orchestrator on the same scoreboard, which is the only honest basis for a renewal or migration decision.

OAuth/OIDC (plain English): standard ways for an agent to prove who it is and what it may do, using your enterprise identity provider instead of long-lived API keys in vendor vaults. **On Monday:** the ITSM gateway accepts tokens from both the CRM agent and the internal ops agent, with different scopes, so you can revoke one without rewiring the other.

Internal tool gateway pattern

Many enterprises will need an internal gateway between agents and systems of record:

Tool gateway, briefly: a single governed front door where agents authenticate, schemas are enforced, budgets apply, and audit events emit before any downstream system is touched. **Architecture case:** agents never hold SAP credentials; they call

```
gateway.company.eu/v1/postings
```

with allowlisted fields and a daily spend cap.

Gateway responsibilities:

Function	Why the board should care
Authenticate agents and users	One revoke point in incidents
Enforce tool allowlists and schemas	Stops prompt-injection tool storms
Emit immutable audit events	Regulator and customer due diligence
Apply budget stop conditions	FinOps visibility before invoice shock
Normalise identity across suite agents	No duplicate CRM + M365 + ITSM credentials
Rate-limit and circuit-break	Prevents cascading failures (OWASP agentic risk)

Gateway cost is platform investment. Compare it to the cost of duplicated credentials across CRM, productivity, and ITSM agents as suite vendors ship agent features through 2026. For a 400-person services firm, three duplicate ITSM integrations often exceed one gateway team-quarter in year-one migration cost alone.

Worked example: Milan-headquartered professional services group

A fictional but realistic composite: 450 staff, Malta and Milan offices, Microsoft 365, Salesforce, and ServiceNow as systems of record.

Month 0: Salesforce ships an agent that wants ServiceNow credentials to “resolve client onboarding blockers.” Microsoft ships a Copilot agent with a similar ITSM plugin. A specialist document agent wants read access to the same contract repository.

Without interoperability discipline: three OAuth apps, three audit formats, three data-processing addenda, and no single owner who can revoke access when a prompt-injection test fails.

With discipline: 1. **Inventory tools in a registry.** Record every business function agents may call with owner, data class, and permitted orchestrators, so new suite agents cannot duplicate credentials silently. The registry is the artefact that turns “three vendors want the same credential” into a governed decision rather than three separate approvals that nobody reconciles. Without it, the security team discovers the duplication only at the next audit. 2. **Route writes through an internal gateway.** ServiceNow and ERP mutations pass a shared schema and audit stream, so CRM and productivity agents do not each hold their own write path. The gateway enforces allowlists and rate limits that a per-vendor integration cannot, and it gives the CIO one place to apply a kill switch. A bad prompt-injection test revokes one token, not three vendor relationships. 3. **Require weekly trace export to the evaluation harness.** Vendor UI screenshots are not evidence; replayable JSON traces feed red-team cases and FinOps attribution every week. This cadence catches a looping agent before month-end invoice shock, and it gives the eval team real failure cases to harden rather than synthetic ones. The clause costs nothing in the contract and everything in incident response. 4. **Procure gateway endpoint support in contract.** Vendors must call customer-operated gateway URLs where feasible, not only embedded connectors that lock you to their orchestrator runtime. This clause preserves the option to swap orchestrators behind the same credential surface, which is the whole point of owning

the spine. Without it, “included in our stack” becomes an uplift of forty per cent at renewal with no portable tool contract.

Outcome the CIO reports to the board: one credential surface, one audit stream, two orchestrators (suite plus internal) measured on cost per resolved onboarding case, not on demo quality.

Protocol boundaries: MCP, A2A, and what boards should ask

Developer communities will argue acronyms. Boards should ask for outcomes: portable auth, auditable tool calls, replaceable orchestrator, stable schemas for critical tools.

MCP (Model Context Protocol): an open pattern for connecting models to tools and data with documented auth; a model-to-tool wire format, not a security product. **On Monday:** MCP lets you swap Claude Opus-class for GPT-5.x (or DeepSeek-V4 behind the same schema) on the same ticket-creation tool without rewriting the ServiceNow adapter.

A2A (Agent-to-Agent), in one line: a pattern for agents to discover each other and exchange tasks; an agent-to-agent wire format, not a substitute for identity policy. **On the P&L:** your research agent delegates drafting to a policy agent; A2A standardises the handoff message, but your gateway still owns credentials.

Neither MCP nor A2A makes a deployment safe by itself. NIST AI 600-1, OWASP, and platform guidance still require scoped identity, authorisation, structured boundaries, tracing, tool restrictions, and human or policy gates.

Board procurement should ask which boundary a vendor is selling, tools, agents, or both, and which security controls remain the buyer’s responsibility.

Tool registry: the minimum inventory

Tool registry, plainly: a living list of every business function an agent may call, with owner, data class, auth method, and review date; the same hygiene APIs already needed, now agent-shaped. **On the P&L:** before approving another agent SKU, finance sees whether the tool already exists and what it costs per thousand calls.

Registry fields for every tool:

Field	Purpose
Owner	Accountable human, not “IT”
Data class	Public / internal / confidential / regulated
Auth method	OIDC scope, mTLS, break-glass rules
Rate limits and blast radius	What one bad call can break
Last review date	Prevents zombie privileges
Orchestrators permitted	Explicit allowlist
Evaluation cases	Misuse and injection tests

Agents make it easy to add tools casually. Each tool is a new privilege surface. OWASP’s 2026 agentic list names tool misuse and excessive agency among top risks. A registry is how security teams sleep.

Multi-vendor reality and bargaining position

Expect Salesforce-like, Microsoft-like, and specialist agents to coexist. Your identity provider and data contracts become the spine. Treat “our agent platform” claims sceptically if they require duplicating system credentials everywhere.

For buyers, this is also a bargaining position. If a vendor cannot expose audit logs, respect enterprise auth, or explain exit assistance, the convenience of one suite may become tomorrow’s migration cost.

Regulation (EU) 2024/1689 adds deployer duties when AI influences consequential decisions; agents that write to systems of record are closer to that line than chat assistants.

Eurostat reports 20.0 per cent of EU enterprises with 10 or more employees used AI in 2025. Interoperability investment is easier while agent estates are still forming than after credentials have fossilised inside three suites.

Counter-position: one vendor stack is simpler

Another board might argue that standardising on one hyperscaler or suite stack avoids gateway cost and protocol debate. That can be true for the first twelve months. It becomes expensive when:

- **The suite reprices agent SKUs separately from core licences.** Agent features arrive as a new line item at renewal, turning "included in our stack" into an uplift of forty per cent with no portable tool contract. The board approved a platform, not a recurring tax on every workflow that touched a tool.
- **A specialist workflow needs a second orchestrator the suite does not support.** Niche document or industry systems require an internal or specialist agent the suite cannot host, forcing duplicate integrations anyway. The single-stack promise breaks at the first industry-specific tool.
- **An incident requires revoking one agent without disabling the whole tenant.** Without a gateway, pulling credentials for a compromised CRM agent may break unrelated productivity workflows on the same platform. One bad agent should not take down the email of every employee.
- **Gartner's cancellation forecast materialises and migration is required.** Exit without documented tool schemas and audit export becomes a six-month rewrite instead of an orchestrator swap behind the same gateway. The cancellation forecast is not abstract; it names weak cost and risk controls as drivers.

Require documented exit assistance and the ability to invoke critical tools through an internal gateway where justified, even in a "single stack" strategy.

CIO decision criteria (board-ready checklist)

Use this at architecture review or renewal:

Criterion	Pass	Fail
Tool API documented with version policy	OpenAPI or equivalent published	"Contact professional services"
Enterprise OIDC/ OAuth supported	Scopes map to your IdP	Shared service account only
Full audit export	JSON/CEF stream you own	Vendor UI only
Budget controls	Per-tool or per-agent caps	Unlimited tool loop
Data residency in contract	EU region named	"Best efforts"
No training on customer data by default	Opt-in only	Broad licence grant
Agent vs assistant labelling	Clear in SKU	Marketing blur
Exit assistance	Data export + connector docs	Silent EOL

Fail three or more on a consequential tool (payments, HR, customer commitments) and defer agent write access until the gateway pattern is in place.

Scenario: suite agent versus specialist agent

Your productivity suite ships an agent. Your CRM ships an agent. Both want credentials to the same ticketing system. Without a gateway and tool registry, you grant twice, log twice, and revoke never.

Interoperability discipline is how you survive the 2026 “agents everywhere” product cycle Gartner describes, without turning identity into spaghetti.

FinOps tie-in: tool calls are money events

FinOps Foundation surveys show nearly all practitioners now manage AI spend. An agent that loops on tools without attribution can burn budget in an afternoon. Interoperability includes stable tool IDs in cost exports so the CFO can see which workflow drove a spike, not only which orchestrator brand.

Finance reading: month-end shows ServiceNow write costs up 300%; trace export proves one CRM agent retried failed closes in a loop; gateway rate limit prevents recurrence.

Interoperability procurement clause list

Copy into RFP and renewal addenda:

1. **Documented tool API with backward-compatible versioning.** Publish OpenAPI or equivalent with a stated deprecation policy so gateway adapters survive vendor minor releases without emergency rewrites. A vendor that cannot commit to a versioning policy is signalling that breaking changes will arrive on your schedule, not theirs. This clause is the difference between a maintenance ticket and a platform re-engineering project.

2. **Enterprise OAuth or OIDC with scoped delegation.** Agents authenticate through the buyer’s IdP with revocable scopes, not long-lived service accounts buried in vendor-managed vaults. Scoped delegation means a compromised agent loses one scope, not the keys to the tenant. The IdP becomes the spine; the vendor becomes a consumer of identity, not a custodian of it.

3. **Full audit export to customer-owned storage.** Immutable JSON or CEF streams land in storage the buyer controls, enabling incident review and FinOps attribution without vendor UI gatekeeping. The export must be machine-readable and contractually guaranteed, not a “premium feature” added at renewal. Without it, the first regulator question after an incident cannot be answered from your own systems.

4. **Budget and rate controls at tool level.** Per-tool caps stop runaway retry loops from burning budget in an afternoon, especially on write tools that agents may invoke repeatedly. The control belongs at the tool, not the orchestrator, because a loop in one agent should not drain the budget of every workflow that shares the tool. Pair the cap with an alert so the on-call operator learns before finance does.

5. **Data residency and subprocessors listed.** Contract names EU regions and subprocessors per data class so legal can map agent tool calls to GDPR and sector obligations. “Best efforts” is not a residency commitment; a named region is. The list must update on renewal, not silently mid-term, because subprocessor changes are data-processing events under GDPR.

6. **No training on customer data by default.** Training use is opt-in only with explicit legal approval, not buried in a broad licence grant on the order form. A broad licence grant lets the vendor improve its model on your confidential prompts and then sell that improvement back to you. Opt-in with legal sign-off makes the data boundary a conscious decision, not a footnote.

7. **Exit assistance with schema docs and bulk export.** Vendor commits to connector documentation, bulk trace export, and a transition window if the SKU ends or reprices materially. The transition window matters because migration behind a gateway takes weeks, not

days, and the vendor holds the schemas. Without this clause, end-of-life is your problem on the vendor's schedule.

8. **Clear agent versus assistant labelling in product and contract.** SKUs distinguish draft-only assistants from tool-using agents so procurement and security apply the right controls. The label drives the control set: assistants need logging, agents need allowlists, budgets, and gates. A SKU that blurs the line lets the vendor sell a drafting tool at agent prices while the buyer under-controls a tool-using one.

9. **Evaluation harness access for replay traces and holdout cases.** Buyer can replay production traces and run red-team packs outside the vendor console, not only view aggregate dashboards. Replay is how a red-team reproduces a failure and how an eval team hardens against the next one. A vendor that refuses this clause is asking you to trust a dashboard you cannot verify.

10. **Support for customer-operated gateway endpoints where feasible.** Vendor agents must call buyer-hosted tool URLs when justified, preserving orchestrator optionality behind one credential surface. The clause keeps the gateway as the spine and the orchestrator as a replaceable tenant. Without it, the vendor's runtime becomes the only path to the tool, and exit cost is whatever the renewal calendar says it is.

Closing position

Interoperability is not a developer argument. It is how boards keep optionality while agents become a default checkbox in enterprise software.

Own tools and identity. Rent orchestrators carefully. Measure workflows, not agent brands.

When Gartner forecasts both agent proliferation and mass cancellation, the firms that survive are those that treated tool access as infrastructure, not as a feature bundled into whichever suite sold first.

References

1. Gartner, "Gartner Predicts 40% of Enterprise Apps Will Feature Task-Specific AI Agents by 2026, Up from Less Than 5% in 2025," press release, 26 August 2025. https://www.gartner.com/en/newsroom/press-releases/2025-08-26-gartner-predicts-40-percent-of-enterprise-apps-will-feature-task-specific-ai-agents-by-2026-up-from-less-than-5-percent-in-2025

2. Gartner, "Gartner Predicts Over 40% of Agentic AI Projects Will Be Canceled by End of 2027," press release, 25 June 2025. https://www.gartner.com/en/newsroom/press-releases/2025-06-25-gartner-predicts-over-40-percent-of-agentic-ai-projects-will-be-canceled-by-end-of-2027

3. Deloitte AI Institute, "The State of AI in the Enterprise: The Untapped Edge" (2026 edition); survey of 3,235 leaders, Aug-Sep 2025. https://www.deloitte.com/us/en/what-we-do/capabilities/applied-artificial-intelligence/content/state-of-ai-in-the-enterprise.html

4. OWASP GenAI Security Project, "OWASP Top 10 for Agentic Applications for 2026," 9 December 2025. https://genai.owasp.org/resource/owasp-top-10-for-agentic-applications-for-2026/

5. FinOps Foundation, "State of FinOps 2026 Report." https://www.finops.org/insights/state-of-finops/

6. McKinsey & Company / QuantumBlack, "The State of AI: Global Survey 2025." https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai

7. Regulation (EU) 2024/1689 of the European Parliament and of the Council (EU AI Act). https://eur-lex.europa.eu/eli/reg/2024/1689/oj

8. Model Context Protocol (MCP) specification and authorization guidance. https://modelcontextprotocol.io/specification/2025-03-26

9. NIST, "Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile," NIST AI 600-1, July 2024. https://nvlpubs.nist.gov/nistpubs/ai/nist.ai.600-1.pdf

10. Eurostat, "20% of EU enterprises use AI technologies," 11 December 2025 (20.0% in 2025; 13.5% in 2024). https://ec.europa.eu/eurostat/en/web/products-eurostat-news/w/ddn-20251211-2

11. Reuters, "Over 40% of agentic AI projects will be scrapped by 2027, Gartner says," 25 June 2025. https://www.reuters.com/business/over-40-agentic-ai-projects-will-be-scrapped-by-2027-gartner-says-2025-06-25/

Published by AIMonger . https://aimonger.com . Malta . Europe