● Ainary

AR-001-v2 Confidence: 70%

State of AI Agent Trust 2026

Only 11% of enterprises have agentic AI in production. Over 40% of projects will be canceled. The EU AI Act arrives in six months. Should you invest in agent trust infrastructure now — or wait?

February 2026

v2.0

Florian Ziesche · Ainary Ventures

> *"Gartner predicts over 40% of agentic AI projects will be canceled by the end of 2027, due to escalating costs, unclear business value or inadequate risk controls."*
>
> — Gartner, June 2025

CONTENTS

FOUNDATION

ANALYSIS

ACTION

1 How to Read This Report

This report uses a structured classification system for every claim. Each claim carries one of four badges:

BADGE	MEANING	EXAMPLE
[E] Evidenced	Backed by external citation from Source Log	8.6% of enterprises have AI agents in production (TechRepublic, n=120K)
[I] Interpretation	Reasoned inference from evidence; logic explained	Agent trust infrastructure is a distinct market segment (synthesized from multiple sources)
[J] Judgment	Recommendation; trade-offs and value assumptions explained	Enterprises should invest now using a "buy + extend" approach
[A] Assumption	Stated but not proven	EU AI Act enforcement will not be delayed beyond August 2026

Confidence levels (High / Medium / Low) reflect the number and quality of independent sources. This report was produced using a **multi-agent research pipeline** with structured source logging, claim ledgering, cross-validation, and adversarial review. Full methodology is in the Transparency Note (Section 13).

2 Executive Summary

Enterprises should invest in AI agent trust infrastructure now, using a "buy + extend" approach — adopt an existing governance platform and customize it — rather than building from scratch or waiting for the market to mature.

- **Agentic AI adoption is 8–14% in production** [E] — TechRepublic reports 8.6% (n=120K); Deloitte reports 11–14%. G2 claims 57%, but uses a broader definition that includes simple automation. The range is wide, but even the low estimate shows acceleration from near-zero in 2024.[1][3][10]
- **Over 40% of agentic AI projects will be canceled by end of 2027** [E] — Gartner attributes cancellations to "escalating costs, unclear business value or inadequate risk controls."[1][10]
- **AI agent security incidents doubled from 2024 to 2025** [E] — prompt injection caused 35% of incidents; some led to $100K+ in real losses.[6]
- **EU AI Act enforcement begins August 2, 2026** [E] — tiered penalties: up to €35M/7% for prohibited practices, €15M/3% for high-risk AI violations, €7.5M/1.5% for other non-compliance.[9]
- **73% of enterprises report a disconnect between agent ambitions and deployment reality**[16] [E] — and only 1 in 5 has a mature governance model for autonomous AI agents[1].
- **Partnerships are 2x more likely to reach production** [E] — Deloitte finds pilots built through strategic partnerships reach full deployment at twice the rate of internal builds.[1]

Keywords: *AI Agent Trust, Agentic AI Governance, EU AI Act Compliance, ISO 42001, NIST AI RMF, Agent Security, Buy vs Build*

3 Methodology

This report synthesizes 16 sources across industry surveys (Deloitte n=unspecified, McKinsey n=1,993, G2 n=5 vendors, TechRepublic n=120K), market research (Precedence Research), regulatory analysis (EU AI Act via LegalNodes), vendor documentation (IBM, Obsidian Security, Vectra AI), security incident reports (Adversa AI), standards bodies (NIST, ISO), and one internal cross-reference (AR-001). The multi-agent research pipeline included structured source logging, claim ledgering (20 claims classified before writing), cross-validation (10 claims verified), contradiction registration (3 contradictions + 2 found in validation), and gap analysis. Confidence is 70% overall — justified by strong regulatory evidence, consistent directional signals on adoption, but wide variance in exact adoption numbers and limited independent TCO data for trust infrastructure specifically.

Limitations: Several market sizing claims rely on a single research firm (Precedence Research). The 50-agent collapse incident is a single researcher's account, not independently verified. G2's 57% adoption figure likely reflects definitional inflation. No independent study compares build vs buy for agent trust infrastructure specifically — the 2x partnership finding is extrapolated from general agent deployment data.

4 Market Reality: Adoption, Acceleration, and the Wide Range Problem 70%

This section answers: How many enterprises actually have AI agents in production, and how fast is adoption growing?

Agentic AI production deployment sits between 8% and 14% of enterprises — not the 57% some surveys claim. [E] TechRepublic's survey of 120,000+ enterprise respondents found only 8.6% have AI agents deployed in production, with 14% developing agents in pilot form[3]. Deloitte's Tech Trends 2026 reports 14% of enterprises have agentic AI solutions ready for deployment and 11% actively in production[1]. Camunda's 2026 report independently corroborates the ~11% figure: only 11% of use cases made it into full production last year. These three sources — using different methodologies and samples — converge on the 8–14% range for genuinely agentic systems.

[E] G2's August 2025 survey reports 57% of companies have AI agents in production[2]. This figure is 4–7x higher than other sources. The likely explanation is definitional: G2 surveyed customers of 5 enterprise AI vendors and used a broad definition of "AI agents" that includes simple automation and chatbots. The 57% figure should not be compared directly to the 8–14% range for agentic AI systems with autonomous decision-making.*

* Footnote: G2's methodology surveyed users of 5 specific enterprise vendors, creating self-selection bias. Their definition of "AI agents" encompasses any AI-powered automation, not the narrower "agentic AI" definition used by Deloitte and TechRepublic.

[E] **Acceleration is unanimous across sources.** Gartner predicts 33% of enterprise software applications will include agentic AI by 2028, up from less than 1% in 2024[10] — a 33x increase in four years. The agentic AI market is valued at $7.55B in 2025 and projected to reach $10.86B in 2026, growing at 43.84% CAGR[4]. McKinsey's 2025 survey found 88% of organizations deploy AI in at least one function, though only 6% are classified as "winning" with deep integration[11]. Early adopters report meaningful impact within 3–6 months[2], suggesting the value realization timeline is compressing for well-governed deployments.

[E] **But 73% of enterprises report a disconnect between agent ambitions and deployment reality**[16], and Gartner predicts over 40% of agentic AI projects will be canceled by end of 2027 due to "escalating costs, unclear business value or inadequate risk controls"[1][10].

CLAIM

[I] The wide adoption range (8–57%) is itself evidence that the market lacks a shared definition of "AI agent." This definitional ambiguity makes vendor claims unreliable and complicates investment decisions. Logic: if three credible surveys produce a 7x range, the surveys are measuring different things.

WHAT WOULD INVALIDATE THIS?

If a standardized definition of "AI agent" emerges and a survey using that definition shows production adoption above 30% by mid-2026, the "still early" framing weakens.

SO WHAT?

For the CTO deciding on trust infrastructure investment: the market is early (8–14% production) but accelerating fast (33x projected by 2028). The 40% cancellation prediction means enterprises that deploy without governance will disproportionately fail. Early investment in trust infrastructure is a competitive differentiator, not premature.

5 The Trust Gap: Incidents, Failures, and Missing Governance 65%

This section answers: What goes wrong when enterprises deploy AI agents without trust infrastructure?

AI agent security incidents doubled from 2024 to 2025, and the most dangerous failures involved agentic AI systems — not simple chatbots. [E] Adversa AI's 2025 incident report found that 35% of all real-world AI security incidents were caused by prompt injection, and some incidents led to $100K+ in real losses. Agentic AI caused the most dangerous failures: crypto thefts, API abuses, and supply chain attacks[6]. (Caveat: Adversa AI is a vendor with commercial incentive to emphasize threats. Their counting methodology is behind a download gate.)

Reported Incidents

[E] **Supply chain attack on the OpenAI plugin ecosystem:** Credentials from 47 enterprises were reportedly harvested through a compromised agent integration[7]. (Source: WebProNews aggregation of multiple reports.)

[E] **50-agent ML system collapsed in 6 minutes:** A single compromised agent reportedly triggered a catastrophic cascade that brought down an entire multi-agent system within six minutes[7]. *[Reported — single source: Akshay Mittal, PhD researcher, writing in InfoWorld. Not an independently verified enterprise incident report. Cited as illustrative of cascading failure risk, not as a documented enterprise case study.]*

[E] **Researchers deployed 44 AI agents and faced 1.8 million attacks** and 62,000 breaches[7]. This demonstrates the attack surface scale when agents are internet-connected.

The Governance Gap

[E] **Only 1 in 5 enterprises has a mature governance model for autonomous AI agents** (Deloitte State of AI 2026)[1]. This means 80% of enterprises deploying agents are doing so without formal governance structures.

[E] **Trust remains the central concern:** G2 reports accuracy, explainability, and security as the top concerns cited by enterprises deploying AI agents[2]. Yet concern has not translated into governance investment — the AI governance tooling market is only $309M in 2025[5], a fraction of the $7.55B agentic AI market.

[I] **The governance market is 24x smaller than the agentic AI market it's supposed to govern.** Calculation: $7.55B (agentic AI) / $309M (governance) = 24.4x. This ratio suggests a structural underinvestment in governance relative to deployment. Logic: governance tooling should grow proportionally with deployment, but is lagging by more than an order of magnitude.

WHAT WOULD INVALIDATE THIS?

If AI agent incident rates plateau or decline in 2026 without new governance tooling — suggesting existing security practices are sufficient — the urgency argument weakens.

SO WHAT?

The trust gap is quantifiable: incidents are doubling, 80% lack governance, and the governance market is 24x smaller than the deployment market. Enterprises deploying agents without trust infrastructure are accepting unquantified risk. The question is not whether incidents will occur, but how costly they will be when they do.

6 Regulatory Pressure: EU AI Act and the August 2026 Deadline 85%

This section answers: What regulatory deadlines and penalties apply to enterprises deploying AI agents?

EU AI Act enforcement begins August 2, 2026 — in less than six months — with a tiered penalty structure that reaches €35 million or 7% of global revenue for the most serious violations.

[E] The EU AI Act uses risk-based classification: unacceptable (banned), high-risk (strict requirements), limited risk (transparency obligations), and minimal risk (no requirements)[9]. The penalty structure is tiered:

Exhibit 1: EU AI Act Penalty Tiers

VIOLATION TYPE	MAXIMUM PENALTY	APPLIES TO
Prohibited AI practices (Article 5)	€35M or 7% of global annual revenue	Social scoring, real-time biometric surveillance, manipulative AI
High-risk AI system violations	€15M or 3% of global annual revenue	Non-compliance with high-risk requirements (Annex III categories)
Other non-compliance	€7.5M or 1.5% of global annual revenue	Providing incorrect information, failing transparency obligations

Source: EU AI Act Article 99 via LegalNodes [9]. Penalty is whichever is higher: fixed amount or percentage of revenue.

[E] Organizations must continuously monitor AI systems, report incidents, and cooperate with authorities. High-risk AI systems in Annex III categories may have extended compliance timelines up to December 2027, but the core enforcement date is August 2, 2026[9].

[E] **Most enterprise AI agent deployments will fall under high-risk or limited-risk categories** — not the prohibited tier. This means the operative penalty for most enterprises is €15M/3%, not €35M/7%. The €35M/7% figure applies specifically to prohibited practices such as social scoring or manipulative AI[9].

[I] **The regulatory deadline is a fixed forcing function, independent of market maturity.** Unlike market dynamics that can be waited out, August 2026 arrives regardless of whether an enterprise has 5 agents or 500. Enterprises that plan to deploy agents in the EU have a fixed deadline for compliance infrastructure. Logic: regulatory deadlines do not adjust to adoption curves.

WHAT WOULD INVALIDATE THIS?

If the EU delays enforcement beyond August 2026, or if implementing guidance (Code of Practice expected June 2026) significantly narrows the scope of AI agent coverage, the urgency decreases.

SO WHAT?

For any enterprise operating in or selling to the EU: compliance infrastructure is not optional after August 2026. The €15M/3% penalty for high-risk AI violations — the tier most relevant to enterprise AI agents — makes non-compliance a material financial risk. The "wait" option has a hard deadline.

7 Trust Frameworks: ISO 42001, NIST AI RMF, and the Vendor Landscape 70%

This section answers: What frameworks and tools exist for implementing agent trust infrastructure?

Two primary trust frameworks have emerged — ISO 42001 (certifiable, audit-ready) and NIST AI RMF (voluntary, faster to implement) — and early adopters are pursuing both simultaneously.

[E] **ISO 42001** provides a certifiable AI management system standard. **NIST AI RMF** offers a voluntary framework organized around four functions: GOVERN, MAP, MEASURE, MANAGE[12]. These frameworks are complementary, not competing: ISO 42001 provides the audit structure; NIST AI RMF provides the operational process.

[E] **Dayforce achieved both ISO 42001 certification and NIST AI RMF attestation in February 2026** — one of the first enterprise HCM vendors to obtain both[13]. This demonstrates that dual-framework compliance is practical and achievable for mid-to-large enterprises.

The Vendor Landscape

[E] **IBM watsonx.governance 2.3.x** (December 2025) added agent inventory management, behavior monitoring, decision evaluation, and hallucination detection[14]. The AI governance tooling market is valued at $309M (2025), projected to grow to $419M (2026) at 35.74% CAGR[5].

[E] **The technical architecture for agent trust is converging** around five components[7][8]:

- **Agent identity and discovery:** Agent Name Service (ANS), Google A2A Protocol, Anthropic MCP, IBM ACP, Linux Foundation AAIF
- **Monitoring and observability:** Real-time behavioral monitoring, anomaly detection

- **Policy enforcement:** Open Policy Agent (OPA), role-based access controls
- **Governance platforms:** IBM watsonx.governance, emerging vendors
- **Audit logging:** Tamper-proof records for compliance and incident response

[I] No single authoritative reference defines "agent trust infrastructure" as a unified category. The five-component architecture above is synthesized from multiple vendor and research sources. Logic: the convergence of multiple vendors building similar components suggests a category is forming, even if it lacks a formal name.

WHAT WOULD INVALIDATE THIS?

If major cloud providers (AWS, Azure, GCP) release comprehensive agent governance services that make standalone governance platforms unnecessary, the "buy a governance platform" recommendation shifts to "activate your cloud provider's built-in governance."

SO WHAT?

The frameworks exist (ISO 42001 + NIST AI RMF). The vendor tooling is emerging (IBM, others). Dual certification is achievable (Dayforce proves it). The question is no longer "is this possible?" but "how fast can we implement it?"

8 Build vs Buy vs Wait 60%

This section answers: Should enterprises build trust infrastructure internally, buy a platform, or wait for the market to mature?

Partnerships and platform-based approaches are twice as likely to reach production as internal builds — and the regulatory deadline eliminates the "wait" option for EU-operating enterprises.

[E] Deloitte's Tech Trends 2026 found that **pilots built through strategic partnerships are 2x more likely to reach full deployment** compared to those built internally, with employee usage rates nearly double for externally built tools[1].

Exhibit 2: Build vs Buy vs Wait Trade-Off Matrix

OPTION	TIME TO PRODUCTION	COST (EST.)	SUCCESS RATE	RISK
Build internally	12–18 months	$1–3M	Baseline (1x)	Misses Aug 2026 deadline; talent-intensive
Buy + extend	3–6 months	$200K–$1M/yr	2x baseline	Vendor lock-in; v1.0 tooling may be immature
Wait	N/A	$0 now; unknown later	N/A	Non-compliant after Aug 2026; deploying agents without governance

Source: Author synthesis. Deployment success rate from Deloitte [1]. Cost estimates from AR-001 [16] [Internal — not independent]. Timeline estimates based on vendor claims and Dayforce dual-certification timeline [14].

[J] **The "buy + extend" approach is recommended** for most enterprises. Trade-offs: buying v1.0 governance tooling carries risk of vendor lock-in and feature immaturity, but the 2x deployment success rate and 3–6 month implementation timeline make it viable before the August 2026 EU AI Act deadline. Building internally delivers more customization but takes 12–18 months — too long for the regulatory timeline. Waiting is the cheapest option today but the most expensive option after August 2026.

[I] The 2x success rate finding applies to agent deployment generally, not trust infrastructure specifically. No study directly compares build vs buy for governance tooling. However, the logic transfers: trust infrastructure requires the same integration expertise, vendor relationships, and operational maturity as agent deployment itself.

WHAT WOULD INVALIDATE THIS?

If enterprise-grade open-source agent governance frameworks emerge (comparable to Kubernetes for container orchestration) before mid-2026, the "buy" recommendation weakens in favor of "adopt open-source + customize."

SO WHAT?

For the CTO: "buy + extend" is the pragmatic choice. Start with a governance platform now, customize for your specific agent portfolio, and iterate as the market matures. The alternative — building from scratch or waiting — either misses the deadline or accepts unquantified compliance risk.

9 Counterargument: "Too Early to Invest"

This section answers: What is the strongest argument against investing in agent trust infrastructure now, and does it hold?

The strongest argument against investing now is that the market is too nascent — only 11% production adoption, immature tooling, and evolving standards mean you're buying v1.0 that will be obsolete in 18 months.

The "too early" case has real evidence behind it:

- [E] **Only 8–14% of enterprises have agents in production**[1][3] — you can't govern what barely exists
- [E] **Only 1 in 5 enterprises has mature AI agent governance** (Deloitte 2026) [1] — the market hasn't defined "good governance" yet
- [E] **The governance tooling market is only $309M**[5] — indicating limited enterprise demand so far
- [E] **ISO 42001 adoption is minimal** — Dayforce is "one of the first" to certify[14], suggesting few have done so
- [I] **v1.0 governance platforms will evolve rapidly** — buying now means re-buying or migrating in 12–18 months

[J] **The counterargument is partially valid but does not change the recommendation.** Three factors override the "wait" logic:

1. **The regulatory deadline is fixed.** [E] August 2, 2026 does not adjust to market maturity. Enterprises operating in the EU need compliance infrastructure regardless of how many agents they run today[9].
2. **Trust infrastructure compounds.** [I] Unlike agent tooling that can be swapped, governance processes — audit trails, incident reporting, risk classification procedures — accumulate institutional knowledge over time. Starting 6 months earlier means 6 months of organizational learning that cannot be back-filled. Logic: compliance is a process capability, not a product purchase.

3. **The cost of failure exceeds the cost of premature investment.** [I] Governance platform costs ($200K–$1M/year) are recoverable if the market pivots. A single high-risk AI violation (up to €15M/3%) is not. The asymmetry favors early investment even if the tooling evolves. Calculation: $1M annual governance cost vs. up to €15M maximum high-risk penalty = up to 15x cost differential.

[J] **The honest answer:** for enterprises with zero agents in production and no near-term deployment plans, waiting 6–12 months is defensible. For enterprises with agents in production or in pilot — the 8–14% that Deloitte and TechRepublic identify — investing now is the risk-adjusted correct decision. The "too early" argument applies to the market broadly, but not to enterprises already in the game.

10 Adversarial Self-Review

This section answers: What would four hostile critics say about this report?

Perspective 1: CFO / Budget Skeptic

"You're asking me to spend $200K–$1M/year on governance for agents that represent 11% of our tech stack. Show me the ROI."

Valid critique. This report lacks independent TCO or ROI data for agent trust infrastructure specifically. The cost asymmetry argument (governance cost vs. penalty risk) is logically sound but unproven in practice. **Mitigation:** Recommend starting with the lowest-cost tier — framework alignment (NIST AI RMF is free) and basic audit logging — before committing to a full governance platform.

Perspective 2: Vendor / Competitor

"This report recommends 'buy + extend' but doesn't evaluate any specific vendor. It's a category recommendation without product validation."

Valid critique. Only IBM watsonx.governance is mentioned by name. No Forrester Wave or Gartner Magic Quadrant was accessible for this report. **Mitigation:** The recommendation is intentionally vendor-agnostic. Enterprises should run their own vendor evaluation; this report provides the decision framework, not the vendor shortlist.

Perspective 3: Academic / Methodologist

"The adoption data is a mess. You acknowledge a 7x range (8–57%) and then pick 11% as your baseline because it fits your narrative. That's not rigorous."

Partially valid. The 8–14% range is supported by three converging sources (TechRepublic, Deloitte, Camunda) vs. one outlier (G2). The resolution is methodologically defensible: prefer the sources with larger samples and stricter definitions. But the report should be clearer that "11%" is a chosen baseline, not a consensus figure.

Perspective 4: "Twitter Critic" / Hostile Reader

"Another AI governance report that says 'invest now!' — brought to you by someone who would benefit from enterprises buying AI governance services. The conflict of interest is obvious."

Acknowledged. Ainary Ventures' business includes AI strategy advisory. This report's conclusion (invest in trust infrastructure) aligns with Ainary's commercial interests. **Mitigation:** Every claim is sourced, classified, and carries an invalidation condition. The counterargument section presents the case for waiting. The reader can evaluate the evidence independently. The report includes one internal source [16] clearly labeled "[Internal — not independent]."

11 Recommendations 65%

This section answers: What should a CTO / VP Engineering do with this information?

Decision Criteria

[J] Invest now if your enterprise meets ANY of these criteria:

- AI agents in production or pilot (you're in the 8–14%)
- Operating in or selling to the EU (August 2026 deadline)
- Industry classified as high-risk under EU AI Act Annex III (healthcare, financial services, HR, critical infrastructure)
- Prior AI-related incident or near-miss

[J] Consider waiting 6–12 months if ALL of these are true:

- No agents in production or planned for 2026
- No EU operations or customers
- Low-risk industry classification

Phased Implementation Plan

Phase 1 (Month 1–2): Foundation

1. Inventory all AI agents — production, pilot, and shadow deployments
2. Classify each agent by EU AI Act risk tier
3. Align to NIST AI RMF (free, voluntary, fast to implement)
4. Establish basic audit logging for all agent actions

Phase 2 (Month 3–4): Platform

1. Evaluate governance platforms (IBM watsonx.governance, emerging vendors)
2. Deploy monitoring and behavioral observation on highest-risk agents
3. Begin ISO 42001 gap assessment (if certification is a goal)

Phase 3 (Month 5–6): Compliance

1. Complete EU AI Act compliance documentation for high-risk systems
2. Establish incident reporting procedures
3. Test human oversight effectiveness (measure actual review rates, not policy existence)
4. Target: operational before August 2, 2026

12 Predictions BETA

These predictions will be scored publicly at 12 months (February 2027). Scoring methodology at ainaryventures.com/predictions.

PREDICTION	TIMELINE	CONFIDENCE
A single AI agent failure causes >$100M in damages (financial, legal, or reputational)	By Feb 2027	50%
At least one enterprise receives an EU AI Act penalty specifically related to AI agent deployment	By Dec 2027	60%
The AI governance tooling market exceeds $450M in annual revenue (reflecting potential acceleration beyond Precedence Research's $419M baseline estimate)	By end 2026	55%

These predictions are falsifiable and testable. In 12 months, results will be published: what was right, wrong, and missed.

13 Transparency Note

This section discloses methodology, known limitations, and confidence calibration.

Overall Confidence

70% — justified by: strong regulatory evidence (EU AI Act dates/penalties verified across multiple legal sources), consistent directional signals on adoption (3 sources converge on 8–14%), but wide variance in exact adoption numbers and no independent TCO data for trust infrastructure.

Sources

16 sources: 3 primary research (Precedence Research x2, McKinsey), 1 standard (NIST), 2 official (Gartner, Dayforce/GlobeNewsWire), 7 reputable secondary (Deloitte, G2, TechRepublic, Adversa AI, WebProNews, LegalNodes, McKinsey), 2 vendor blogs (Obsidian Security, Vectra AI), 1 internal (AR-001).

Strongest Evidence

EU AI Act enforcement date and penalty tiers (verified across LegalNodes, artificialintelligenceact.eu, multiple law firms). Gartner >40% cancellation prediction (confirmed via Gartner's own press release, June 2025).

Weakest Point

No independent study compares build vs buy for agent trust infrastructure specifically — the 2x partnership finding is extrapolated from general agent deployment data (Deloitte). Market sizing relies on a single firm (Precedence Research). The 50-agent collapse is a single researcher's account.

What Would Invalidate

If EU AI Act enforcement is delayed >6 months, or if agent deployment success rates reach >30% without dedicated governance, or if major cloud providers bundle governance into standard offerings (eliminating "buy" decisions).

Contradictions

3 registered: (1) Adoption range 8.6–57% (definitional, resolved by leading with 8–14% for agentic specifically). (2) Gartner 33% by 2028 vs 40% by 2026 (different reports/definitions, preferred primary Gartner source). (3)

	Rapid adoption AND >40% cancellation (not contradictory — both can be true, reinforces trust infrastructure thesis).
Methodology	Multi-agent research pipeline (A+ Pipeline v2.0): Research Agent (source log + claim ledger), Validation Agent (cross-validation + gap check), Writer Agent (synthesis per template), with structured QA. 20 claims classified before writing. 10 claims cross-validated. Not a systematic literature review — a targeted synthesis for decision support.
Limitations	Several sources are vendor-produced (Adversa AI, Obsidian Security, Vectra AI) with commercial incentives. TechRepublic data could not be deep-verified (Cloudflare block). The Camunda and Deloitte governance findings are cited from secondary references, not primary report access. One internal source (AR-001) is labeled but not independent.
System Disclosure	This report was created with a multi-agent research system. Florian Ziesche directed the research; the AI system executed the pipeline.

14 Claim Register

20 load-bearing claims, classified before writing. Top 5 include invalidation conditions.

Exhibit 3: Claim Register

#	CLAIM	VALUE	SOURCE	TYPE	CONFIDENCE	USED IN
C1	Agentic AI production deployment rate	8–14%	[1][2][3][10][11]	[E]	Med	§4
C2	Agent adoption growth trajectory	33x by 2028	[10][1][2]	[E]	High	§4
C3	Agentic AI market size (2025)	$7.55B	[4]	[E]	Med	§4,5
C4	AI governance market size (2025)	$309M	[5][15]	[E]	Med	§5,7
C5	>40% agentic AI projects canceled by 2027	>40%	[1][10]	[E]	High	§4
C6	Early adopters see ROI in 3–6 months; only 6% "winning"	3–6mo / 6%	[2][11]	[E]	Med	§4
C7	AI agent security incidents doubled 2024–2025	2x	[6][7]	[E]	Med	§5
C8	50-agent collapse, 47-enterprise	Reported	[7]	[E]*	Med	§5

	supply chain attack					
C9	Primary attack vectors for AI agents	5 vectors	[8][6][7]	[E]	High	§5
C10	Trust is #1 enterprise concern for AI agents	#1	[2]	[I]	Med	§5
C11	Two primary frameworks: ISO 42001 + NIST AI RMF	2 frameworks	[12][13][8]	[E]	High	§7
C12	Agent trust infrastructure has 5 technical components	5 components	[7][8][14]	[I]	Med	§7
C13	Partnerships 2x more likely to reach production	2x	[1]	[E]	Med	§8
C14	Governance vendor landscape maturing rapidly	IBM Dec 2025	[14][5]	[E]	Med	§7
C15	"Wait" option carries increasing risk	Aug 2026	[9][1][10]	[J]	Med	§8,9
C16	EU AI Act enforcement Aug 2026,	€35M/7% max	[9][12]	[E]	High	§6

	tiered penalties					
C17	High-risk AI compliance required by Aug 2026, some Annex III by Dec 2027	Aug 2026	[9]	[E]	High	§6
C18	Trust infrastructure cost ($200K–$2M) vs failure cost ($5K–$100M+)	1–2 OoM gap	[15][9][7]	[I]	Med	§8,9
C19	Enterprise should invest NOW (GO)	GO	[9][1][2][10][6][7][5][14]	[J]	Med	§2,11
C20	Recommended approach: "buy + extend"	Buy+extend	[1][14][13]	[J]	Med	§8,11

**Single-source, not independently verified*

Top 5 Claims — Invalidation Conditions

1. **C1 (8–14% adoption):** Invalidated if a standardized-definition survey shows >30% production adoption by mid-2026
2. **C5 (>40% canceled):** Invalidated if Gartner revises this prediction downward, or if cancellation rate is below 25% at year-end 2027
3. **C16 (EU AI Act Aug 2026):** Invalidated if the EU delays enforcement by >6 months
4. **C19 (GO decision):** Invalidated if agent failure costs decline significantly (e.g., insurance products make losses insurable at low premium) or if cloud

providers bundle governance at no incremental cost

5. **C13 (2x partnership success):** Invalidated if independent research shows comparable or higher success rates for internal builds in the trust infrastructure domain specifically

15 References

[1] Deloitte. (2025). "The Agentic Reality Check: Preparing for a Silicon-Based Workforce." *Tech Trends 2026*. deloitte.com. Accessed 2026-02-15.

[2] G2. (2025). "Enterprise AI Agents Report: Industry Outlook for 2026." learn.g2.com. Accessed 2026-02-15.

[3] TechRepublic. (2026). "AI Adoption Trends in the Enterprise 2026." techrepublic.com. Accessed 2026-02-15.

[4] Precedence Research. (2025). "Agentic AI Market Size to Hit USD 199.05 Billion by 2034." precedenceresearch.com. Accessed 2026-02-15.

[5] Precedence Research. (2025). "AI Governance Market Size, Share and Trends 2025 to 2034." precedenceresearch.com. Accessed 2026-02-15.

[6] Adversa AI. (2025). "Top AI Security Incidents of 2025 Revealed." adversa.ai. Accessed 2026-02-15.

[7] WebProNews. (2026). "AI Agents' Trust Reckoning: One Hack Fells 50, Exposing Urgent Need for Digital Identity Backbone." webpronews.com. Accessed 2026-02-15.

[8] Obsidian Security. (2026). "The 2025 AI Agent Security Landscape: Players, Trends, and Risks." obsidiansecurity.com. Accessed 2026-02-15.

[9] LegalNodes. (2026). "EU AI Act 2026 Updates: Compliance Requirements and Business Risks." legalnodes.com. Accessed 2026-02-15.

[10] Gartner. (2025). "Intelligent Agents in AI." gartner.com. Accessed 2026-02-15.

[11] McKinsey & Company. (2025). "The State of AI: Global Survey 2025." mckinsey.com. Accessed 2026-02-15.

[12] NIST. (2025). "AI Risk Management Framework." nist.gov. Accessed 2026-02-15.

[13] GlobeNewsWire / Dayforce. (2026). "Dayforce Achieves ISO 42001 Certification and NIST AI RMF Attestation." globenewswire.com. Accessed 2026-02-15.

[14] Vectra AI. (2026). "AI Governance Tools: Selection and Security Guide for 2026." vectra.ai. Accessed 2026-02-15.

[15] Ainary Research. (2026). "State of AI Agent Trust 2026." AR-001. [Internal — not independent]

[16] Camunda. (2026). "2026 State of Agentic Orchestration and Automation." camunda.com. Accessed 2026-02-15.

Cite as: Ainary Research (2026). "State of AI Agent Trust 2026." AR-001-v2.

About the Author

Florian Ziesche is the founder of Ainary Ventures, where AI does 80% of the research and humans do the 20% that matters. Before Ainary, he was CEO of 36ZERO Vision and advised startups and SMEs on AI strategy and due diligence. His conviction: HUMAN × AI = LEVERAGE. This report is the proof.

ainaryventures.com

● Ainary

AI Strategy · Published Research · Daily Intelligence

Contact · Feedback

ainaryventures.com
florian@ainaryventures.com