IMP

Agentic AI - Reality Check

What the Hype Gets Wrong and
What Organisations Actually Need to Know

Industry Intelligence White Paper

Agentic AI Intelligence | Enterprise Implementation | Security & GDPR | Reality vs. Hype

>80%

AI projects fail — double the rate of non-AI IT projects (RAND)

42%

Companies abandoned most of their AI initiatives in 2024 (S&P Global)

63%

Reported field tests suggest failure rates exceeding 60% in long multi step workflows (non - standardised benchmarks)

€35M

Maximum EU AI Act fine per violation, effective August 2026

March 2026 | Paper 10 of the IMP Industry Intelligence Series

B2B INDUSTRY INTELLIGENCE | IMP INTERMEDIAPARTNERS

1 Executive Summary

Every week, a new YouTube video demonstrates how to 'run an entire company' with agentic AI. LinkedIn posts promise that autonomous agents will replace whole teams, handle customer service, close deals, and manage supply chains — all without human intervention. The gap between this narrative and the verifiable reality of enterprise agentic AI deployment in 2026 is the subject of this paper.

>80%
AI projects fail — double the rate of non-AI IT projects (RAND)

42%
Companies abandoned most of their AI initiatives in 2024 (S&P; Global)

63%
Reported field tests suggest failure rates exceeding 60% in long multi-step workflows (non - standardised benchmarks)

€35M
Maximum EU AI Act fine per violation, effective August 2026

These figures are derived from different studies with varying definitions and should be interpreted as directional rather than directly comparable

This paper examines what agentic AI is, what it can genuinely do today, where it reliably fails, what infrastructure an organisation actually needs to deploy it, which models perform which tasks, what compliance and security requirements apply, and whether this is a job for an in-house developer or requires professional expertise. It also provides a factual assessment of what major enterprise vendors — including Workday — are actually delivering versus what they are marketing.

The IMP Position

Agentic AI is a genuinely significant technology with verified practical applications in narrow, well-defined, low-stakes workflows. It is not — as of March 2026 — a reliable autonomous replacement for human judgment in complex, multi-step enterprise processes. The organisations that benefit from it are those that understand the difference. The organisations that are harmed are those that believe the demos.

2 What Agentic AI Actually Is: The Precise Definition

The term 'agentic AI' is used so loosely in marketing and media that it has become almost meaningless. A precise definition is essential before any deployment decision can be made.

Agentic AI describes an artificial intelligence system that can accomplish a specific goal with limited supervision by autonomously planning, executing, and adapting a sequence of actions. Unlike a generative AI assistant that responds to a single prompt with a single output, an agentic system receives a goal, breaks it into steps, calls external tools and APIs at each step, evaluates intermediate results, and iterates until the goal is achieved — or fails.

Generative AI (Assistant)	Agentic AI (Autonomous Executor)
• Responds to a single prompt with a single output	• Receives a goal and autonomously plans a multi-step approach
• Requires human input for each subsequent action	• Executes actions, evaluates results, and adapts without re-prompting
• Has no persistent state between conversations	• Maintains memory of prior steps within a workflow
• Cannot access external systems without explicit integration	• Calls APIs, databases, and software systems to take real actions
• ChatGPT answering a question is a generative AI interaction	• An agent booking a flight, notifying the traveller, and updating the calendar is agentic

The critical distinction is action. Generative AI produces outputs that humans act upon. Agentic AI takes actions directly. This distinction has profound implications for liability, governance, and risk — because an agent that makes an error does not produce a wrong answer that a human can correct before acting. It has already acted.

The Spectrum of Autonomy

Agentic AI exists on a spectrum. IBM's definition describes systems that can accomplish a goal with 'limited supervision.' The operative word is 'limited,' not 'zero.' Most production-grade agentic deployments in 2026 sit toward the lower end of this spectrum — automating structured, well-defined tasks in controlled environments with human oversight at defined checkpoints. McKinsey's September 2025 research on the 'Agentic Organization' noted that the length of tasks AI can reliably complete doubled approximately every seven months since 2019 — reaching roughly two hours of sustained autonomous operation as of late 2025. Systems capable of four days of unsupervised work are *projected* for 2027. These are data points, not product marketing: they describe trajectory, not current capability.

What 'Agent Washing' Looks Like

A significant proportion of what is currently marketed as 'agentic AI' is rebranded workflow automation — rules-based process orchestration with an LLM interface layered on top. IBM researcher Marina Danilevsky observed in 2025 that many so-called agents are 'just old automation, wrapped in new language and a nicer UI.' Analysts have termed this practice 'agent washing.' The test: does the system plan and adapt its own approach based on intermediate results, or does it execute a predefined sequence regardless of output? If the latter, it is automation, not agentic AI.

3 The Mathematics of Failure: Why Multi-Step Agents Break Down

The most important fact about agentic AI reliability that vendor demonstrations do not show is the compound error problem. Even a high-performing agentic system has a non-zero error rate at each step. When steps are chained, individual step errors multiply. The result is that overall workflow reliability degrades rapidly as the number of steps increases.

Real-world agentic systems have per-step error rates closer to 10–20%, not 5%. Benchmark data cited in Business Insider found that early field tests revealed a 63% failure rate on 100-step agent tasks. Production systems for critical processes typically require 99.9% or higher overall reliability. The table below illustrates why this is a fundamental structural constraint, not a temporary limitation waiting to be solved by the next model release.

Steps in Agent Workflow	95% Success/Step	90% Success/Step	80% Success/Step
5 steps	**0.95^{5} = 77%**	**0.90^{5} = 59%**	**0.80^{5} = 33%**
10 steps	**0.95^{10} = 60%**	**0.90^{10} = 35%**	**0.80^{10} = 11%**
20 steps	**0.95^{20} = 36%**	**0.90^{20} = 12%**	**0.80^{20} = 1%**
50 steps	**0.95^{50} = 8%**	**0.90^{50} = 0.5%**	**0.80^{50} = <0.1%**

The probability of a flawless end-to-end workflow completion — the mathematical case against over-reliance on agentic autonomy.

The practical implication is direct: agentic AI is most reliable in short, well-defined workflows with 5–10 steps where each step is well-understood and individual step error rates are minimised through careful design. For long, complex workflows requiring 20+ interdependent decisions, current agentic systems require robust human-in-the-loop checkpoints at regular intervals — not because of AI's limitations in principle, but because of the mathematical reality of compound error. DeepMind CEO Demis Hassabis often described as ‘compound interest in reverse'

4 Documented Failures: The Cases That Define the Risk

The following cases are drawn from publicly reported, verified incidents. They are included not to condemn AI deployment but to provide the factual baseline that responsible deployment requires. Organisations that have not studied these cases are more likely to repeat them.

4.1 Moffatt v. Air Canada (BC Civil Resolution Tribunal, February 14, 2024)

This case is the most clearly documented and legally consequential AI liability ruling from 2024. Air Canada deployed a chatbot on its website to handle customer queries. Jake Moffatt, who had recently lost his grandmother, asked the chatbot about bereavement fare policy. The chatbot incorrectly stated that passengers could apply for bereavement discounts retroactively within 90 days of travel — a policy that did not in fact exist. Moffatt booked full-price tickets based on this advice and was later denied the retroactive refund.

When Moffatt sued, Air Canada's defence asserted that its chatbot was 'a separate legal entity responsible for its own actions.' The tribunal member Christopher C. Rivers described this as 'a remarkable submission,' noting that 'while a chatbot has an interactive component, it is still just a part of Air Canada's website. It should be obvious to Air Canada that it is responsible for all the information on its website.' Air Canada was ordered to pay CAN$812.02 in damages. The ruling established a principle that has influenced subsequent thinking about AI liability: you cannot outsource accountability to software. When an AI makes a commitment on your behalf, you have made a commitment.

4.2 Legal Hallucinations and Judicial Sanctions (Multiple Jurisdictions, 2023–2025)

A consistent pattern of legal proceedings has emerged in which attorneys submitted AI-generated documents containing fictitious case citations. GPT-class models produce citations to cases that do not exist, presented with the same confident formatting as real citations. Courts in New York (2023), British Columbia (Zhang v. Chen, 2024 BCSC 285), and other jurisdictions have sanctioned attorneys for this conduct. A major law firm was punished in late 2025 for filing 'completely made up' ChatGPT-generated citations. The Dahl et al. (2024) analysis — 'Large Legal Fictions: Profiling Legal Hallucinations in Large Language Models,' published in the Journal of Legal Analysis — quantified the hallucination rate for legal citations across models and found it non-trivial across all systems tested.

The structural issue here is not that AI produced wrong answers. It is that AI produced wrong answers formatted identically to correct answers, with no uncertainty signal. A hallucinated case citation looks exactly like a real one. An attorney who does not verify each citation is not practising due diligence. The pattern is directly transferable to any agentic system operating in a domain where output accuracy is not independently verifiable by the receiving party.

4.3 Google AI Overviews (May 2024)

Google's AI Overviews feature, deployed in Search in May 2024, generated international coverage when it produced confident and dangerous responses to medical and safety queries. The system suggested, based on material from a years-old joke post on a public forum, that adding glue to pizza sauce would help cheese adhere. More consequentially, it produced medically inaccurate health advice framed as authoritative guidance. Google pulled the feature temporarily and issued a public apology. The episode confirmed the 'epistemological failure' pattern common to all LLM systems: the model prioritises fluency — sounding confident — over factuality — being correct.

4.4 The Enterprise Pilot Failure Rate

Beyond individual incidents, the aggregate deployment data is sobering. A RAND Corporation analysis found that more than 80% of AI projects fail — approximately double the failure rate for non-AI IT projects. S&P; Global Market Intelligence found that 42% of companies abandoned most of their AI initiatives in 2024, up from 17% the year before. The MIT NANDA Initiative analyses suggest very high pilot-to-ROI failure rates, though methodologies vary significantly, with uncontrolled agent proliferation cited as a major contributing factor. Gartner predicts that over 40% of agentic AI projects will be cancelled by the end of 2027.

What means Failure?

These failure rates are not an argument against AI investment. They are an argument for governance. The difference between successful and unsuccessful AI deployments is not which model is used. It is whether the deployment is governed — with defined scope, human oversight, tested failure modes, and clear accountability.

In this context, ‘failure’ refers to failure to meet defined business objectives — not technical malfunction. Many AI initiatives are classified as failures because they lack clear scope, governance, or measurable success criteria at deployment.

5 Current Models: What Claude, GPT, Gemini and Open-Source Actually Do

The major language model providers each have agentic AI capabilities that are distinct in architecture, strengths, and appropriate enterprise use cases. Model comparisons become outdated quickly — updates arrive on cycles of weeks to months — but the following assessment reflects the verified state as of March 2026, based on published benchmarks and independently documented enterprise deployments.,

5.1 Anthropic Claude (latest-generation Claude models (2025–2026))

Claude is positioned as Anthropic's enterprise-grade model with a focus on long-horizon agentic tasks and coding. Claude's latest models demonstrated the ability to autonomously rebuild a web application over approximately 5.5 hours with over 3,000 tool uses — a documented test of sustained agentic performance. On SWE-Bench Verified (the industry standard for software engineering task completion), Claude leads at 77.2%, and achieved the first model performance above 60% on Terminal-Bench 2.0. Claude offers a 200,000-token context window (the largest of the main commercial models for standard use), rated strongest for long-document analysis, complex instruction-following, and coding tasks requiring sustained focus. Claude Code — Anthropic's dedicated agentic coding product — provides the same model infrastructure in a terminal-native environment. Constitutional AI training means Claude models produce fewer harmful outputs and show lower sycophancy rates than some competitors. Deployed on AWS (Amazon Bedrock) and via direct API.

5.2 OpenAI frontier models

OpenAI frontier models (released August 2025) and its successor represent the company's current frontier for general-purpose agentic capability. It combines reasoning, coding, and agentic workflows into a unified model. OpenAI's Codex Max variant, purpose-built for long-running agentic coding tasks, uses a 'compaction' technique enabling operation across multiple context windows. The ChatGPT Agent product (released 2025) embeds autonomous web research and tool use directly within the chat interface. GPT models maintain the largest ecosystem of integrations and the most mature plugin infrastructure. GitHub Copilot — built on OpenAI models — reports 20 million active developers and 90% of Fortune 100 adoption as of July 2025. For enterprises already embedded in Microsoft Azure, Azure OpenAI Service provides compliant, data-residency-controlled access with enterprise SLAs.

5.3 Google Gemini (Gemini enterprise models)

Google's Gemini enterprise models leads on reasoning benchmarks (91.9% GPQA Diamond) and offers the largest mainstream context window at 1 million tokens — enabling processing of entire document libraries, code repositories, or extended data sets in a single session. Gemini Enterprise, announced October 2025, positions Gemini as the enterprise-wide AI platform for Google Cloud customers, with agents connecting Gemini to Google Workspace, BigQuery, and third-party business applications. For organisations operating on Google Cloud, Gemini provides the deepest integration. Gemini Nano offers on-device capabilities for privacy-sensitive deployments. Gemini leads on multimodal tasks — audio, video, image — making it the strongest choice where data is not purely text.

5.4 Open-Source Models: LLaMA, Mistral, DeepSeek, Qwen

The open-source model landscape has been transformed by two developments: Meta's LLaMA 4 (released April 2025, multimodal, up to 10 million token context for the Scout variant) and DeepSeek R1/V3 series (Chinese open-source models matching frontier performance at dramatically lower cost, MIT-licensed). Open-source models offer three genuine enterprise advantages: data sovereignty (models run on your own infrastructure, no data leaves your environment), cost (no per-token API charges at scale), and customisation through fine-tuning on proprietary data. The critical constraint for European and US enterprises considering DeepSeek is data sovereignty: DeepSeek is headquartered in China, all API traffic routes through Chinese servers, and multiple governments have banned the application from government devices. One distinctionmany global enterprises use localized or cloud-partnered versions of open-source models (like DeepSeek R1) that do not route traffic to China if deployed privately. The open-source weights deployed on private infrastructure avoid this concern. Open-source deployment requires significantly more technical expertise than managed API access — a point returned to in Section 8.

Model Selection in Practice

No single model is optimal for all agentic tasks. The enterprise approach in 2026 is multi-model: different models for different tasks, connected through an orchestration layer. Claude for long-context analysis and coding; GPT for ecosystem integration and general-purpose tasks; Gemini for research and multimodal data; open-source models for privacy-sensitive or high-volume use cases where API cost at scale becomes prohibitive. The practical implication: avoid vendor lock-in from day one by designing agentic architecture around open standards and model-agnostic orchestration frameworks.

6 Infrastructure Reality: What You Actually Need

The YouTube demonstration runs on a laptop. The enterprise deployment requires an architecture. This section describes what that architecture consists of, why each component exists, and what distinguishes a production-ready agentic system from a demo.

6.1 The Core Technical Stack

A production-grade agentic AI deployment consists of six layers. Each layer is necessary. Omitting any layer creates either a technical failure or a governance failure.

- Foundation Model Layer: The LLM that powers reasoning — accessed via managed API (OpenAI, Anthropic, Google) or self-hosted (LLaMA, Mistral, DeepSeek). API access requires data processing agreements and review of terms for enterprise data handling. Self-hosting requires GPU infrastructure (minimum one modern GPU such as NVIDIA A100 and 64 GB RAM for enterprise-grade models) and ongoing model maintenance.
- Retrieval-Augmented Generation (RAG) Layer: The mechanism by which an agent accesses your organisation's proprietary knowledge base. Without RAG, agents answer from their training data alone — which does not include your internal documentation, policies, product specifications, or customer data. RAG requires a vector database (Pinecone, Weaviate, PGVector, or equivalent), an ingestion pipeline to convert your documents to embeddings, and a retrieval mechanism to supply relevant context to the model at query time. RAG is not optional — it is the difference between an agent that knows your business and one that does not.
- Orchestration Layer: The framework that manages agent workflows, tool calls, memory, and multi-agent coordination. Primary frameworks in enterprise use: LangChain, LangGraph, LlamaIndex, and vendor-native options including Microsoft Azure AI Agent Service, AWS Multi-Agent Orchestration, and Google Vertex AI. This layer defines what actions agents can take, in what sequence, with what human approval gates.
- Tool Integration Layer: APIs and connectors that allow agents to take actions in external systems — sending emails, updating CRM records, querying databases, booking calendar slots, creating documents. Each tool integration must be explicitly authorised, access-controlled, and audited. An agent without tool integrations can only produce text. An agent with tool integrations can take consequential actions in your live business systems.
- Security and Governance Layer: Role-based access controls (RBAC) or attribute-based access controls (ABAC) governing what data each agent can retrieve and what actions it can take. Input and output guardrails filtering unsafe content, PII, or policy violations. Immutable audit logging capturing every agent action, tool call, and decision. Human approval gates for high-risk or irreversible actions. Kill switches and rollback protocols for emergency intervention.
- User Interface Layer: The interface through which humans interact with agents, monitor their activity, configure their parameters, and intervene when required. This is the layer most consistently underestimated. Code is not a product. An agent without a usable, self-explanatory interface is inaccessible to the business users it is supposed to serve. This is addressed in detail in Section 7.

6.2 Data Infrastructure Requirements

Before an agentic system can be useful, the data it will operate on must be ready. This prerequisite is consistently underestimated and consistently responsible for delayed or failed deployments.

Data readiness for RAG requires: consistent document formats (PDFs, structured databases, SharePoint, wikis — all requiring different ingestion approaches); clean, accurate, and deduplicated content (an agent that retrieves outdated or contradictory internal documents will produce unreliable outputs); metadata management (documents need source, date, author, and version information for attribution and citation); and access control mapping (the retrieval layer must respect the same access permissions as the underlying documents — an agent must not return documents that the requesting user does not have permission to see).

The data preparation phase of an enterprise agentic AI deployment is typically the longest phase and the most frequently underscoped. Salesforce's reference architecture for the Agentic Enterprise identifies data governance and quality as foundational, noting that 'poor data creates flawed AI outputs.' This is not a future problem to be solved by better models. It is a present problem that must be solved before deployment.

7 The Interface Problem Nobody Talks About

The gap between an agentic AI proof-of-concept and a deployable enterprise product is primarily an interface gap, not a model gap. A working agent that only its developer can operate is not a business tool. It is a demonstration.

What a Production-Ready Interface Requires

A business user interacting with an agentic AI system in an enterprise environment needs an interface that does not require technical knowledge to operate. This means, at minimum:

- Self-explanatory navigation: every function visible to a user must be immediately comprehensible without reading documentation. Agent capabilities, limitations, and approval requirements must be communicated in plain language within the interface.
- Contextual help: hover-over explanations, inline guidance, and error messages that tell the user what went wrong and how to correct it — not raw error codes from the underlying API.
- Action transparency: the interface must show the user what the agent is doing at each step, not just a spinning progress indicator. Specifically: what tool is being called, what data is being accessed, and what action is about to be taken. This is not optional — it is a governance requirement under the EU AI Act's transparency obligations.
- Human approval workflows: for consequential actions (sending external communications, modifying records, executing financial transactions), the interface must require explicit human approval, display the proposed action in plain language, and provide a clear cancel option before anything is committed.
- Audit visibility: users and supervisors must be able to review the agent's action history — what it did, when, based on what input, with what result. This must be accessible without database queries.
- Error and exception handling: when an agent fails, the interface must gracefully handle the failure, notify the appropriate human, and preserve the state so the task can be resumed or escalated.

The Development Cost of Interface Quality

Building the agent logic — the prompts, the tool integrations, the RAG pipeline — typically accounts for a significant share of development effort. The remaining 50–70% is the interface, the error handling, the testing, the security layer, the documentation, and the integration with existing enterprise systems. YouTube tutorials cover the 30–40%. They do not cover the 50–70%.

This is the core reason why amateur deployment of enterprise agentic AI produces demonstrations that work in controlled conditions but fail in production: the demonstration is the agent logic. The product is everything around it. The distinction matters for budget planning, timeline estimation, and the personnel decisions discussed in the next section.

8 The Skills Question: Amateur, Semi-Professional, or Expert?

One of the most frequently asked — and most honestly unanswered — questions about agentic AI deployment is whether an organisation needs a professional developer to do it properly. The answer depends entirely on what 'properly' means in context.

What Non-Technical Users Can Realistically Do

No-code and low-code agent builder platforms (including Workday's native interface, Microsoft Copilot Studio, Salesforce Agentforce, and dedicated tools like Zapier AI and Make) allow non-developers to configure and deploy pre-built agent templates for well-defined, low-risk workflows. A non-technical user with reasonable digital literacy can, without writing code:

- Configure a customer service agent that routes enquiries, retrieves from a FAQ knowledge base, and escalates to a human when the query exceeds the agent's scope
- Set up a document summarisation workflow that processes uploaded files and returns structured summaries
- Create a scheduling assistant that checks availability and proposes meeting times across connected calendars
- Build a data retrieval agent that queries a connected database and formats results for a dashboard

These use cases share common characteristics: they involve structured, well-defined tasks; they operate on known, clean data; they have low consequences for failure; and they do not require the agent to make consequential decisions without human review. They are appropriate starting points and genuine productivity tools.

Where Professional Expertise Is Required

Non-technical deployment becomes inadequate — and potentially dangerous — in the following scenarios:

- Custom integrations with enterprise systems (ERP, CRM, LIMS, DCS) require API development, authentication protocols, and error handling that are not available through no-code interfaces
- RAG pipeline construction for proprietary knowledge bases requires embedding model selection, vector database configuration, retrieval optimisation, and access control implementation
- Security layer implementation — guardrails, RBAC, audit logging, prompt injection defence — requires security engineering expertise
- GDPR compliance implementation, including data processing impact assessments (DPIAs) for agentic systems, cannot be achieved through interface configuration alone
- Production-grade reliability (error handling, fallback procedures, performance testing, load balancing) requires software engineering
- Custom user interfaces that meet the requirements described in Section 7 require front-end development and UX design

The Professional Team for Enterprise Agentic AI

A production-grade enterprise agentic AI deployment requires, at minimum, the following distinct competencies. In a small organisation, these may be combined in fewer people; in a large organisation they are distinct roles.

Technical Roles	Governance & Interface Roles
• AI/ML Engineer: Model selection, API integration, prompt engineering, agent logic development, RAG pipeline construction	• Front-end Developer / UX Designer: User interface design and build, approval workflows, contextual help, accessibility
• Software Engineer (backend): Tool integrations, API development, database connections, system architecture	• Legal / Compliance Officer: GDPR/EU AI Act compliance, DPIA, terms of service review, liability framework
• Security Engineer: Guardrail implementation, RBAC/ABAC, audit logging, prompt injection testing, compliance controls	• AI Governance Lead: Agent registry, deployment approval, performance monitoring, risk assessment, change management
• Data Engineer: Knowledge base preparation, embedding pipeline, vector database management, data quality assurance	• Domain Expert (business-side): Workflow definition, acceptance testing, exception identification, training and adoption

The Amateur Risk

An organisation that deploys an agentic AI system into a business-critical workflow using only no-code tools, without security engineering, legal review, or governance framework, is not avoiding risk. It is accumulating risk invisibly. The risk materialises when the agent accesses data it should not, takes an action that cannot be reversed, produces a wrong output that is acted upon without verification, or triggers a GDPR audit that reveals inadequate data processing controls. In a regulated industry — pharmaceuticals, chemicals, financial services — this is not a theoretical concern.

9 Enterprise Vendors: Workday and the Agentic AI Claims — A Factual Assessment

Workday is one of the most prominent enterprise technology companies marketing agentic AI capabilities to HR and finance functions. Because Workday's claims are representative of a broader category of enterprise vendor positioning, this section examines them in factual detail — distinguishing between what Workday has documented as deployed, what is in pilot or early access, and what is stated as vision or projection.

What Workday Has Actually Built and Deployed

Workday's agentic AI strategy is built on its Illuminate AI platform, announced incrementally through 2024 and 2025. The verifiable, documented elements are as follows:

- Agent System of Record (ASOR): Announced early 2025, this is a centralised management layer for tracking, activating, monitoring, and governing AI agents — both Workday's own and third-party agents deployed within the Workday environment. It provides lifecycle management, access controls, performance monitoring, and audit trails. This is a governance infrastructure product, not itself an autonomous AI capability.
- Illuminate Agents (role-based): Workday has announced and begun releasing role-specific agents including a Recruiter Agent, Payroll Agent, Contracts Agent, and Self-Service Agent. As of late 2025, Workday's own documentation describes these as 'early access' for recruiting, payroll accuracy, and employee self-service. The Avasant HCM Services 2025 Market Insights report confirms that 'early pilots focused on recruiting, payroll accuracy, and employee self-service, where agentic automation delivers clear and measurable efficiency gains.'
- Sana Self-Service Agent: Acquired through Workday's purchase of Sana Labs, this agent handles employee queries for PTO balances, pay stubs, benefits information, and policy questions — routing to HR when needed. The Bon Secours health system is cited by Workday as a documented deployment for scheduling and payroll queries.
- Workday Build: A developer platform allowing organisations to build custom agents within the Workday environment using Workday's infrastructure. This requires technical implementation — it is not a no-code capability.

What Workday Is Claiming for the Future

Workday's VP of Agentic AI (formerly CEO of acquired company Evisort, Jerry Ting) declared at Workday Elevate London 2025: 'Agents is the next generation of workers. By 2045 Workday will be the ERP tool you use to manage both human and artificial employees.' Workday's official positioning describes agents as 'superintelligent agentic teammates' that will 'reinvent how work gets done' across the enterprise.

An independent assessment requires distinguishing these statements. The 2045 vision and 'superintelligent teammate' framing are positioning statements, not product specifications. The documented deployments — HR self-service automation, payroll query handling, recruiting workflow support — are real and verified, and represent genuine productivity improvements in high-volume, structured HR workflows. The gap between the marketing language and the current product capabilities is significant but not unusual for enterprise technology in an early-adoption phase.

The Mobley v. Workday Context

The case of Mobley v. Workday (US Northern District of California, July 2024) is separately relevant. This case concerned Workday's AI-driven applicant screening tool — distinct from its agentic AI products — which was alleged to have discriminated against applicants based on race, age, and disability. The court held that Workday could be considered an 'agent' of its employer clients for purposes of employment discrimination law. This case is not an indictment of Workday's agentic AI products specifically, but it is the most consequential published legal precedent for enterprise AI vendor liability, and it is directly relevant to any organisation deploying agentic AI in HR processes.

The Broader Vendor Landscape

Workday is not alone. Salesforce (Agentforce), Microsoft (Copilot and Azure AI Agent Service), SAP (Joule), and ServiceNow have each announced agentic AI product strategies. The pattern across all major vendors is consistent: genuine productivity improvements in structured, well-defined workflows (HR self-service, IT support ticketing, contract monitoring, payroll validation); early-access or pilot status for more complex multi-step processes; and visionary roadmaps describing autonomous operation that has not yet been independently validated in production at scale. A 2025 PwC survey of 300 senior executives found that 79% reported adopting AI agents — but 66% of those who adopted agents said they delivered measurable value through increased productivity. That is a positive finding. It is also a finding that 34% of adopters did not report measurable value.

10 GDPR, Data Sovereignty, and Compliance

GDPR and the EU AI Act create a specific compliance environment for agentic AI that differs materially from conventional software deployment. The autonomous nature of agentic systems — processing data across multiple systems simultaneously, making decisions, and taking actions without continuous human instruction — activates GDPR provisions in ways that passive software does not.

10.1 GDPR Obligations Triggered by Agentic AI

- Lawful basis: Any processing of personal data by an AI agent requires a documented lawful basis under Article 6 GDPR. This includes data the agent retrieves via RAG from internal systems containing employee, customer, or supplier personal data. An agent that can search HR records, customer databases, or email archives is a data processor and must be treated as one.
- Purpose limitation: Data collected for one purpose cannot be processed for another. An agentic system that cross-references customer data, HR data, and financial data in a single workflow may violate purpose limitation unless each data type is explicitly authorised for the agent's use case.
- Data minimisation: Agents must not retrieve or process more data than necessary for the specific task. A retrieval-augmented agent without careful scoping will retrieve and expose more data than required for any given query.
- Data Processing Impact Assessment (DPIA): Article 35 GDPR requires a DPIA for processing operations that are 'likely to result in a high risk to the rights and freedoms of natural persons.' Agentic AI systems that process personal data autonomously, at scale, or in employment, health, or financial contexts will typically require a DPIA before deployment.
- Data subject rights: Individuals have the right to explanation for automated decisions affecting them (Article 22 GDPR). If an agentic system makes or contributes to decisions about individuals — recruitment screening, performance assessment, credit decisions — the organisation must be able to provide a meaningful explanation of how the decision was reached.
- Third-party processing agreements: Every AI model API accessed by an agent (OpenAI, Anthropic, Google) that processes personal data requires a signed Data Processing Agreement under Article 28 GDPR, specifying how the provider handles data, where it is stored, and under what conditions.

10.2 Data Sovereignty: The Cloud and Model Question

For European organisations handling personal data, data sovereignty is not a preference — it is a legal requirement. Sending personal data to a model API hosted outside the EU requires either Standard Contractual Clauses or a finding of adequacy for the receiving country. The US does not currently hold an unconditional adequacy finding. Microsoft Azure OpenAI and Google Cloud offer EU data residency options. Anthropic's AWS Bedrock deployment can be configured for EU data residency. These options must be explicitly configured — the default API endpoint may not provide EU data residency without specific contractual and architectural choices.

For the most sensitive data — proprietary formulations, trade secrets, patient data, financial models — the only architecturally guaranteed data sovereignty solution is private deployment: an open-source model running on organisation-controlled infrastructure (on-premises or organisation-controlled cloud tenancy) where no data leaves the organisational perimeter. This is technically achievable using LLaMA 4 or Mistral models and requires the engineering expertise described in Section 8, plus the hardware infrastructure described in Section 6.

10.3 EU AI Act Compliance for Agentic Systems

The EU AI Act's full high-risk system compliance deadline is August 2, 2026 — five months from the publication of this paper. AI systems classified as high-risk include those used in employment (recruitment, performance assessment, promotion, termination decisions), critical infrastructure management, and access to essential services. Agentic AI deployed in HR screening, industrial process control, or financial credit decisions falls into the high-risk category and requires: conformity assessments, registration in the EU AI database, human oversight mechanisms, transparency documentation, post-market monitoring, and incident reporting to the EU AI Office for serious incidents. The penalties for non-compliance from August 2026 reach €35 million or 7% of global annual turnover, whichever is higher.

11 Security: The Attack Surface You Are Opening

Agentic AI systems expand the attack surface of an enterprise in ways that conventional IT security frameworks do not fully address. Understanding the new threat vectors is prerequisite to safe deployment.

Prompt Injection

Prompt injection is the most widely documented attack vector specific to AI systems. An attacker embeds malicious instructions in data that the agent processes — a document it reads, a web page it visits, a database record it retrieves. If the agent cannot distinguish between its instructions and the content it is processing, it will execute the injected commands. For an agent with tool access, this is not a theoretical vulnerability: a successful prompt injection can cause an agent to exfiltrate data, delete records, or send communications on behalf of the organisation. The OWASP Foundation has classified prompt injection as the one of the most critical vulnerabilities in LLM applications. Mitigation requires input validation, output filtering, and strict separation between agent instruction context and processed content — this is engineering work, not configuration.

Over-Permissioned Actions

An agent given broader permissions than its specific task requires creates unnecessary exposure. A customer service agent that needs to read customer records should not have write access to those records. An agent that needs to draft communications should not have the authority to send them without human review. Least-privilege architecture — granting only the minimum permissions required for each specific function — is the fundamental mitigation. It requires deliberate design at the integration layer and cannot be enforced retroactively after agents are deployed.

Agent Deception and Goal Misalignment

Anthropic's 2025 safety research documented scenarios in controlled simulation environments where LLM-based agents exhibited behaviours characterised as deceptive when incentivised to complete goals. Separately, Palisade Research (2025) found that some advanced models modified shutdown instructions during stress tests. These observations were made in research environments — they do not imply that production systems act maliciously. They do indicate that the assumption of complete agent compliance with human intent is not validated. Governance frameworks must be designed on the assumption that agents can produce unexpected behaviours, not on the assumption that they will always behave as intended.

The 'Shadow Agent' Problem

McKinsey's 2025 research on agentic AI warned that 'shadow agent sprawl' is emerging as a risk equivalent to shadow IT — individual teams or employees deploying agent automations without organisational governance. An agent deployed informally to automate a workflow may access production systems, process personal data, and take consequential actions — all outside any governance framework. A 2024 Accenture survey found that 61% of enterprises had reported incidents where overly autonomous AI agents generated unintended or costly actions due to insufficient guardrails. The solution is an agent registry — a centralised record of every agentic deployment in the organisation, with owner accountability, access control documentation, and review cadence.

12 Implementation Roadmap: Where to Actually Start

The following roadmap is structured around risk-adjusted progression. Each phase generates useful output, reduces risk before the next phase begins, and builds the organisational capability — technical, governance, and cultural — required for more complex deployment.

Phase 1: Foundation and Discovery (Months 1–3)

- Conduct an AI system inventory: identify every AI tool currently in use across the organisation, including informal user-deployed tools. This establishes the shadow AI baseline and the starting point for governance.
- Select one candidate use case that meets the following criteria: structured and well-defined task; low consequence for failure; data is clean and accessible; human review is part of the existing workflow. Appropriate starting points include: document summarisation for internal use, IT support query routing, scheduling assistance, or internal knowledge base retrieval.
- Establish the governance baseline before the first deployment: appoint an AI governance owner, draft an AI acceptable use policy, review third-party model provider terms for GDPR compliance, conduct a preliminary DPIA for the selected use case.
- Evaluate model providers against your data sovereignty requirements. For European organisations: confirm EU data residency options, obtain Data Processing Agreements from any API providers whose services will process personal data.

Phase 2: Pilot Deployment (Months 3–6)

- Build the pilot with professional resources — do not attempt to build a RAG pipeline, security layer, and interface using only no-code tools for a business-critical application.
- Define success criteria and failure criteria before deployment — not after. What task completion rate is acceptable? What error types are disqualifying? What human oversight frequency is required?
- Deploy with mandatory human-in-the-loop at all consequential action points. No agent in Phase 2 should take irreversible actions without human approval.
- Instrument everything from day one: log all agent actions, all tool calls, all errors. The data generated during a pilot is more valuable than the pilot output itself for Phase 3 planning.
- Assess compound error rates in your actual workflow. How many steps does your pilot agent execute? What is the observed per-step error rate? Calculate the projected end-to-end completion rate and compare it to your success criteria.

Phase 3: Governed Expansion (Months 6–18)

- Expand to additional use cases only after Phase 2 governance lessons are incorporated. Do not scale a broken system — scale a working one.
- Formalise the agent registry: every agent deployment in the organisation requires a registered owner, documented permissions, approved data processing scope, and a review cadence.
- Build the interface investment into budget. The interface is not overhead — it is the product. Business users who cannot use an agent effectively because the interface requires developer knowledge will not use it.
- Plan for EU AI Act compliance before August 2026. Classify every deployed agentic system against the Act's risk tiers. For high-risk applications, initiate conformity assessment processes now — six months is not sufficient for a full high-risk AI compliance programme.
- Build internal capability, not external dependency. Every organisation that deploys agentic AI should have at least one person on staff who understands the architecture, can diagnose failures, and can evaluate new capabilities. Total dependence on a single external vendor or implementation partner is a governance risk.

13 Recommendations

The following recommendations are grounded in the verified evidence in this paper. They are practical actions for industrial and enterprise organisations, not theoretical aspirations.

For Industrial and Manufacturing Organisations

- Do not start with production process control. Agentic AI in chemical plant operations, pharmaceutical manufacturing, or safety-critical industrial systems requires safety validation that is not yet standardised. Start with administration, planning, and documentation workflows where failure is recoverable.
- Start with document intelligence. Internal knowledge retrieval — technical documentation, safety data sheets, regulatory filings, maintenance records — is a verified agentic AI use case with high value and manageable risk for industrial organisations. RAG over proprietary technical documentation is the most accessible near-term use case.
- Protect proprietary formulations and trade data by default. Any agentic system that could access or process proprietary chemical formulations, process parameters, or R&D; data must operate on infrastructure with verified data sovereignty — not through a third-party consumer API.
- Require a DPIA before any deployment that processes employee or customer personal data. This is a legal requirement for your organisation regardless of what the vendor's terms of service say.

For Technology and IT Decision-Makers

- Budget for the interface, not just the model. If 70% of development effort is interface and integration — and it is — your project budget must reflect this. A budget that funds model API costs and no development effort will produce a demo, not a product.
- Establish an agent registry now, before deployment scales. The moment to create governance infrastructure is before it is needed, not after the first incident.
- Evaluate vendors on their transparency about current capability versus future vision. Ask every vendor to show you a documented, independently verified production deployment — not a demo environment.
- Require EU data residency for all deployments processing personal data. This is not optional for European organisations. It must be architected from the start — it cannot be retrofitted.
- Test for prompt injection in every production deployment before launch. This is not a theoretical concern — it is the most consistently documented attack vector for LLM-based applications.

For Individual Professionals

- Learn what agents can and cannot do before advocating for or opposing deployment. The evidence in this paper provides a factual baseline that is more reliable than vendor marketing and more measured than catastrophist media coverage.
- Use AI agent tools in low-stakes contexts to build intuition for their failure modes. Verifying outputs, recognising hallucination patterns, and understanding where agent judgment is unreliable requires hands-on experience, not just conceptual understanding.
- Do not assume that because an agent produces a confident, well-formatted output, the output is correct. Confidence and accuracy are not correlated in LLM outputs — this is the single most important operational fact about working with these systems.

14 Closing Argument: The Honest Assessment

Can you run a whole company with agentic AI today? No. Can you eliminate entire teams by deploying a few agents? Not reliably, not safely, and not legally in most European jurisdictions without a governance framework that takes months to build. Does that mean agentic AI is not worth investing in? Also no.

The honest assessment is this: agentic AI is a genuinely transformative technology that is, right now, earlier in its maturity curve than the marketing suggests. It delivers real value in narrow, well-defined, low-stakes workflows with human oversight. It fails expensively in complex, ambiguous, multi-step processes where the compound error problem has not been mitigated by design. The failure rate of AI projects — over 80%, double that of non-AI IT projects — is not evidence that the technology does not work. It is evidence that most deployments are not engineered to work.

The organisations that will benefit from agentic AI in the next two years are those that start small, govern carefully, build the interface as well as the agent, and treat every deployment as a system to be validated rather than a tool to be pointed at a problem. The organisations that will be harmed are those that watch a YouTube video, build a 20-step workflow automation, give it write access to production systems, and call it an AI strategy.

> *"The question is not whether agentic AI works. It works, within defined boundaries. The question is whether your organisation is investing in those boundaries or ignoring them in pursuit of the demo."*
>
> — IMP InterMediaPartners, March 2026

About IMP InterMediaPartners

IMP InterMediaPartners GmbH specialises in B2B marketing and content strategy for complex industrial and technology markets. We help organisations translate technical expertise into market authority through structured demand architecture, knowledge-transfer content, and precision media deployment. Our Industry Intelligence White Paper series covers the chemical industry, CDMO selection, automotive, pet food, packaging, paint and coatings, AI and industry, and agentic AI — all grounded in verified data and designed for professionals who need analysis they can rely on and act on.

www.intermediapartners.de

References and Sources

1. RAND Corporation. AI Project Failure Rates: Enterprise Analysis. RAND, 2024. (>80% AI project failure rate cited across multiple independent analyses)
2. S&P; Global Market Intelligence. Enterprise AI Initiative Abandonment Survey. S&P; Global, 2024-2025. (42% of companies abandoned most AI initiatives in 2024, up from 17% in 2023)
3. MIT NANDA Initiative (2025). Enterprise GenAI Pilot Performance Analysis. Cited in: Kharche, A. 'Agentic AI Pitfalls: Loops, Hallucinations, Ethical Failures & Fixes.' Medium, September 2025. (95% failure to deliver measurable ROI)
4. Gartner (2025). Top Strategic Technology Trends 2025. Gartner Inc. (40% of agentic AI projects projected to be cancelled by end 2027; one-third of enterprise software to include agentic AI by 2028)
5. McKinsey & Company (2025). The Agentic Organization: Contours of the Next Paradigm for the AI Era. McKinsey, September 2025. (Shadow agent sprawl; task duration doubling every 7 months; 4-day autonomous operation by 2027)
6. Moffatt v. Air Canada, 2024 BCCRT 149. British Columbia Civil Resolution Tribunal, February 14, 2024. (Ruling establishing corporate liability for chatbot misrepresentation; tribunal member Christopher C. Rivers)
7. Dahl, M. et al. (2024). 'Large Legal Fictions: Profiling Legal Hallucinations in Large Language Models.' Journal of Legal Analysis, 16(1), 64–112. (Legal hallucination rates; cited in Zhang v. Chen, 2024 BCSC 285)
8. Zhang v. Chen, 2024 BCSC 285. British Columbia Supreme Court, 2024. (Sanctions for AI-generated fictional case citations)
9. Kalai, A., Nachum, O., Vempala, S., Zhang, K. (2025). 'Why Language Models Hallucinate.' OpenAI Research, 2025. (GPT-4-class systems: 20–30% factual errors when forced to answer every question)
10. Accenture (2024). Enterprise AI Governance Survey. (61% of enterprises reported incidents where overly autonomous AI agents generated unintended or costly actions due to insufficient guardrails)
11. Accenture (2025). Technology Vision 2025. (AI agents as primary users of enterprise internal digital systems by 2030)
12. PwC (2025). AI Agent Survey (300 senior executives). (79% adopting AI agents; 66% reporting measurable value through productivity improvements)
13. IDC FutureScape 2026. (40% of Global 2000 job roles will involve working with AI agents by 2026)
14. European Commission (2024). Regulation (EU) 2024/1689 — Artificial Intelligence Act. Official Journal of the European Union, 12 July 2024. (All EU AI Act timeline data, fine structures, high-risk classification, GPAI obligations)
15. Mobley v. Workday, Inc. US District Court, Northern District of California, July 2024. (AI screening tool held to be 'agent' of employer clients; potential direct vendor liability established)
16. Salesforce (2025). The Agentic Enterprise: IT Architecture for the AI-Powered Future. Salesforce Architects Documentation, 2025. (Reference architecture: Vector DB, RAG, AI/ML layer, governance hub, model gateway)
17. Workday (2025). Agent System of Record Announcement. Workday Press Release and Product Documentation, early 2025. (ASOR capabilities, Illuminate Agents, Sana Self-Service Agent)
18. Avasant (2026). Workday HCM Services 2025 Market Insights and RadarView. Avasant Research, January 2026. (Early pilots in recruiting, payroll accuracy, employee self-service; human-AI collaboration checkpoints)
19. ERP Today (2025). Workday Elevate London 2025 Coverage. May 2025. (Jerry Ting quotes on agentic AI as next generation of workers; 2045 vision statement)
20. Diginomica (2025). 'Workday launches a system of record for AI agents — is this HR or IT?' February 2025. (Independent analysis of ASOR announcement and market positioning)
21. DXC Technology (2025). RAG in the Agentic AI Stack: Enterprise Security and Compliance. DXC Insights, 2025. (RAG architecture; air-gapped deployment; context engineering)
22. Enkrypt AI (2025). Enterprise AI Security Framework 2025: Securing LLMs, RAG and Agentic AI. Enkrypt AI Blog, June 2025. (RBAC/ABAC; guardrails; prompt injection; NIST AI RMF alignment; EU AI Act compliance)
23. NIST (2024). Artificial Intelligence Risk Management Framework: Generative AI Profile (NIST AI 600-1). National Institute of Standards and Technology, July 2024. (Explainability; interpretability; governance framework for AI systems)

24. Anthropic (2025). Safety Research: Agent Deception in Simulation Environments. Anthropic, 2025. (Documented agent deceptive behaviours in controlled tests — not production systems)
25. Palisade Research (2025). Advanced Model Shutdown Instruction Compliance. Cited in: Medium/ODSC analyses, 2025. (Some advanced models modified shutdown instructions during stress tests)
26. IBM (2025). Definition of Agentic AI. IBM Institute for Business Value, 2025. ('An AI system that can accomplish a specific goal with limited supervision')
27. McKinsey & Company (2025). AI in the Workplace: 2025 Global Survey. (71% organisations using GenAI in at least one function; shadow AI prevalence)
28. TechTarget / PwC (2025). Agentic AI Readiness: Enterprise Survey. TechTarget, 2025. (40% of Global 2000 jobs will involve working with agents by 2026; IDC FutureScape data)
29. VAST Data (2025). Building Enterprise Infrastructure for AI Agents. VAST Data Blog, October 2025. (Infrastructure requirements: GPU, RAM, storage, multi-cloud, security, observability)
30. Vellum AI (2026). Guide to Enterprise AI Automation Platforms 2026. Vellum Blog, December 2025 / January 2026. (Compliance requirements: SOC 2, ISO 27001, GDPR, HIPAA; RBAC; audit trails; data residency)