

ENTERPRISE AI AGENTS DEMAND IS REAL. READINESS IS NOT.

What current evidence says about adoption, value, reliability and the path from agent pilots to production.

**40%**

of large-company respondents report scaling AI agents, up from 27% a year earlier.

McKinsey, 2026 [1]

5%

say their business processes are highly prepared for AI agents.

Deloitte, 2026 [2]

37%

report a positive contribution from AI to enterprise EBIT; only 6% are high performers.

McKinsey, 2026 [1]

68%

of production agents studied execute ten steps or fewer before human intervention.

IBM Research, 2026 [3]

THE CORE FINDING

The market opportunity is not generic agent development. It is redesigning high-value workflows into reliable, governed and measurable agent systems.

EXECUTIVE SUMMARY

Companies need agents - but not everywhere

The evidence supports investment in agents, while warning against starting with autonomy, platforms or hype.

RESEARCH CONCLUSION

Enterprises do have a growing need for agents where work requires context, judgment, tool use and multi-step coordination. The bigger unmet need is not access to an agent framework. It is the ability to select the right workflow, integrate it safely, evaluate it continuously and operate it reliably.

Five things leaders should take from the data

1

Demand is scaling

Large enterprises are moving fastest: 40% of respondents at organizations above \$1B revenue report scaling agents. [1]

2

Readiness trails ambition

Only 5% say their processes are highly prepared; data, trust, governance and integration remain material barriers. [2]

3

Value depends on workflow redesign

Enterprise EBIT impact is still concentrated. High performers redesign work instead of inserting AI into unchanged processes. [1]

4

Production agents are bounded

Most production agents in IBM's study were short-horizon systems with frequent human intervention; reliability was the top challenge. [3]

5

Governance must match autonomy

Controls should increase with the agent's ability to act and the scope of access it receives. [4]

DECISION RULE

Start with the work. Use an agent only when it is the best mechanism for improving a measurable process outcome.

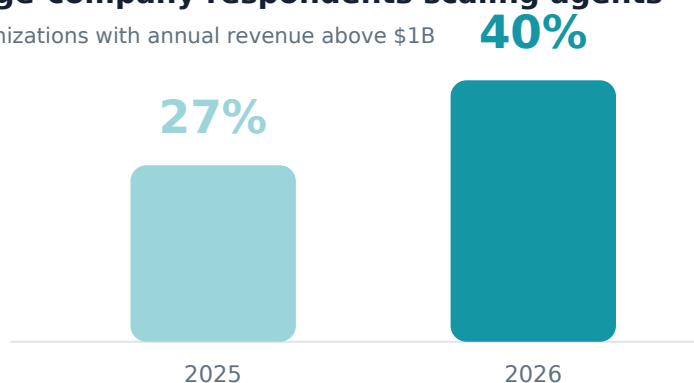
1. DEMAND

Adoption is moving from experiments toward scale

Large organizations are leading, but the strongest signals are concentrated in specific functions and workflow types.

Large-company respondents scaling agents

Organizations with annual revenue above \$1B



Agent scaling rose by 13 percentage points in one year among large organizations, while smaller organizations remained broadly flat at 22%.

Source: McKinsey Global Survey 2026 [1]

Where current scaling is most visible

IT & OPERATIONS

Incident triage, service operations, investigations, ticket handling and controlled action.

KNOWLEDGE WORK

Search, synthesis, policy guidance, reporting and institutional knowledge connected to action.

SOFTWARE ENGINEERING

Coding agents, defect investigation, testing, documentation and release assistance.

INDUSTRY WORKFLOWS

Marketing and sales in consumer sectors; supply chain and manufacturing in advanced industry.

WHAT THE DEMAND SIGNAL DOES NOT PROVE

Adoption is not the same as value. McKinsey reports 60% expect AI investment to rise, but only 37% currently attribute positive EBIT impact to AI. [1]

2. THE GAP

The constraint is organizational, not just technical

Enterprises are testing agents faster than they are preparing processes, data, controls and operating models.



What leaders say is getting in the way



Source: Deloitte survey of 501 U.S. business and IT leaders already piloting agents [2]

The value gap

37%

report a positive contribution from AI to enterprise EBIT.

6%

qualify as AI high performers, reporting at least 5% EBIT impact and significant value.

Source: McKinsey Global Survey 2026 [1]

INTERPRETATION

The evidence does not say agents lack value. It says value is unlikely to appear by layering an agent over an unchanged process without clear ownership, integration and measurement.

3. PRODUCTION REALITY

Useful agents are usually bounded, observable and interruptible

IBM's study of 306 practitioners across 26 domains offers a practical picture of what is actually operating in production.

68%

execute 10 steps or fewer
before human
intervention

70%

use prompting of
off-the-shelf models
rather than weight tuning

74%

depend primarily on
human evaluation

TOP DEVELOPMENT CHALLENGE

Reliability - consistent correct behavior over time. Practitioners address it through system design, not model choice alone. [3]

Match control to autonomy

1

OBSERVE

Read-only retrieval and explanation

2

ADVISE

Recommendations or drafts; human acts

3

ACT WITH APPROVAL

Agent prepares; designated human approves

4

BOUNDED ACTION

Agent acts inside explicit permissions and limits

GOVERNANCE PRINCIPLE

An agent's ability to act and the scope of access it receives should determine the strength of controls. [4]

4. SELECTION

Not every workflow needs an agent

The right solution depends on whether the work needs retrieval, fixed rules, human judgment or adaptive multi-step action.

SEARCH

Find and answer

Use when the main need is grounded retrieval or Q&A.; No tool action is required.

AUTOMATION

Execute fixed rules

Use when steps and decisions are deterministic, stable and fully specified.

COPilot

Assist a person

Use when a human remains the operator and AI drafts, summarizes or recommends.

AGENT

Own a bounded objective

Use when the system must interpret context, choose next steps, use tools, handle exceptions and escalate.

A strong candidate workflow has most of these traits



Frequent enough to justify integration and operating cost



Multi-step work that crosses data, tools or teams



Judgment or exception handling that fixed rules cannot cover



A clear objective and measurable business outcome



Boundaries that can be expressed as permissions and policies



A safe escalation path when confidence is low

A USEFUL ANTI-HYPE TEST

If an assistant or deterministic workflow can deliver the outcome with lower cost and risk, do not use an agent. Gartner explicitly recommends agents for decisions, automation for routine flows and assistants for simple retrieval. [5]

5. PROOF IN PRACTICE

Three workflow patterns from Applied AI Consulting

These are AAIC case studies and target operating models. Targets are not presented as measured production outcomes.

CUSTOMER REPORTING**4 systems | 2 formats | 3 validation roles**

Collect operational evidence, plan the report, draft it, validate data/editorial/production quality, route review through Teams and preserve recovery state.

Pattern: agent orchestration around evidence, validation and human review.

AAIC case study [8]

TELECOM INVOICE RECONCILIATION**6 provider families | 7 deterministic parsers**

Keep extraction and exact matching deterministic. Use agent reasoning for exception investigation with read-only billing data, preserved evidence and fail-closed controls.

Pattern: deterministic core plus bounded agent judgment.

AAIC case study [9]

MULTI-TENANT M365 SERVICE DESK**8 operations | 6-8 min baseline to 1-2 min review target**

Resolve the correct customer and tenant, obtain explicit approval, execute one controlled Graph action, verify the end state and write audit-ready ServiceNow notes.

Pattern: one ticket, one tenant, one action, verified and auditable.

AAIC case study [10]

WHAT THESE EXAMPLES HAVE IN COMMON

They begin with a workflow boundary, preserve evidence, separate deterministic work from model judgment, include human control and verify the result before closure.

6. FROM IDEA TO PRODUCTION

A practical agent opportunity blueprint

Before choosing a model or framework, map the work and define how value, risk and reliability will be managed.

01

WORKFLOW

Trigger, objective, current effort, handoffs, exceptions and outcome owner.

02

DECISIONS

Which judgments need AI? Which decisions must remain deterministic or human-owned?

03

CONTEXT

Approved knowledge, data sources, freshness, tenant boundaries and evidence requirements.

04

TOOLS

Read and write actions, permissions, credentials, rate limits and rollback paths.

05

CONTROL

Autonomy level, approvals, stop conditions, audit trail and escalation rules.

06

EVALUATION

Test set, quality threshold, tool-call checks, regression coverage and production feedback.

07

ECONOMICS

Cycle time, throughput, error/rework, service quality, operating cost and value owner.

RECOMMENDED STARTING POINT

Choose one high-friction workflow. Classify it as non-agentic, agent-assisted, approval-gated or bounded autonomous. Build the smallest safe version that can produce measurable evidence.

CONCLUSION

The need is real. The buying criterion has changed.

Agent-building capability is increasingly available. Production readiness remains scarce.

START WITH THE WORK.

The strongest enterprise offer is not 'we can build an agent.' It is the ability to identify the right workflow, design the operating boundary, connect the systems, test the behavior, govern the actions and own the path to production.

What this means for enterprises

- Invest in workflow discovery and redesign before platform selection.
- Use autonomy as a risk decision, not as a maturity badge.
- Require evaluation, observability and auditability from the first pilot.
- Measure the full process outcome, not only model accuracy or token cost.

Selected sources

- [1] McKinsey & Company (25 Aug 2026). The state of AI in 2026: On the road to ROI.
- [2] Deloitte (Aug 2026). AI Agents are Only the Beginning: the AI readiness gap.
- [3] IBM Research (23 Apr 2026). Measuring Agents in Production.
- [4] Gartner (26 May 2026). Applying uniform governance across AI agents will lead to failure.
- [5] Gartner (25 Jun 2025). Over 40% of agentic AI projects will be canceled by end of 2027.
- [6] McKinsey & Company (14 Sep 2026). Stacking the odds: A blueprint for successfully scaling agentic AI.
- [7] Deloitte (21 Jan 2026). The State of AI in the Enterprise 2026.

Applied AI Consulting case studies

- [8] Customer Reporting Operations with LangSmith Fleets (22 Jul 2026).
- [9] Telecom Supplier Invoice Reconciliation with LangSmith Fleets (21 Jul 2026).
- [10] Multi-Tenant M365 Service Desk Automation with LangSmith Fleets (21 Jul 2026).

ABOUT APPLIED AI CONSULTING

AAIC designs, builds and operationalizes enterprise AI products and agentic workflows across business and IT operations - with integration, evaluation, governance and human control built in. appliedaiconsulting.com