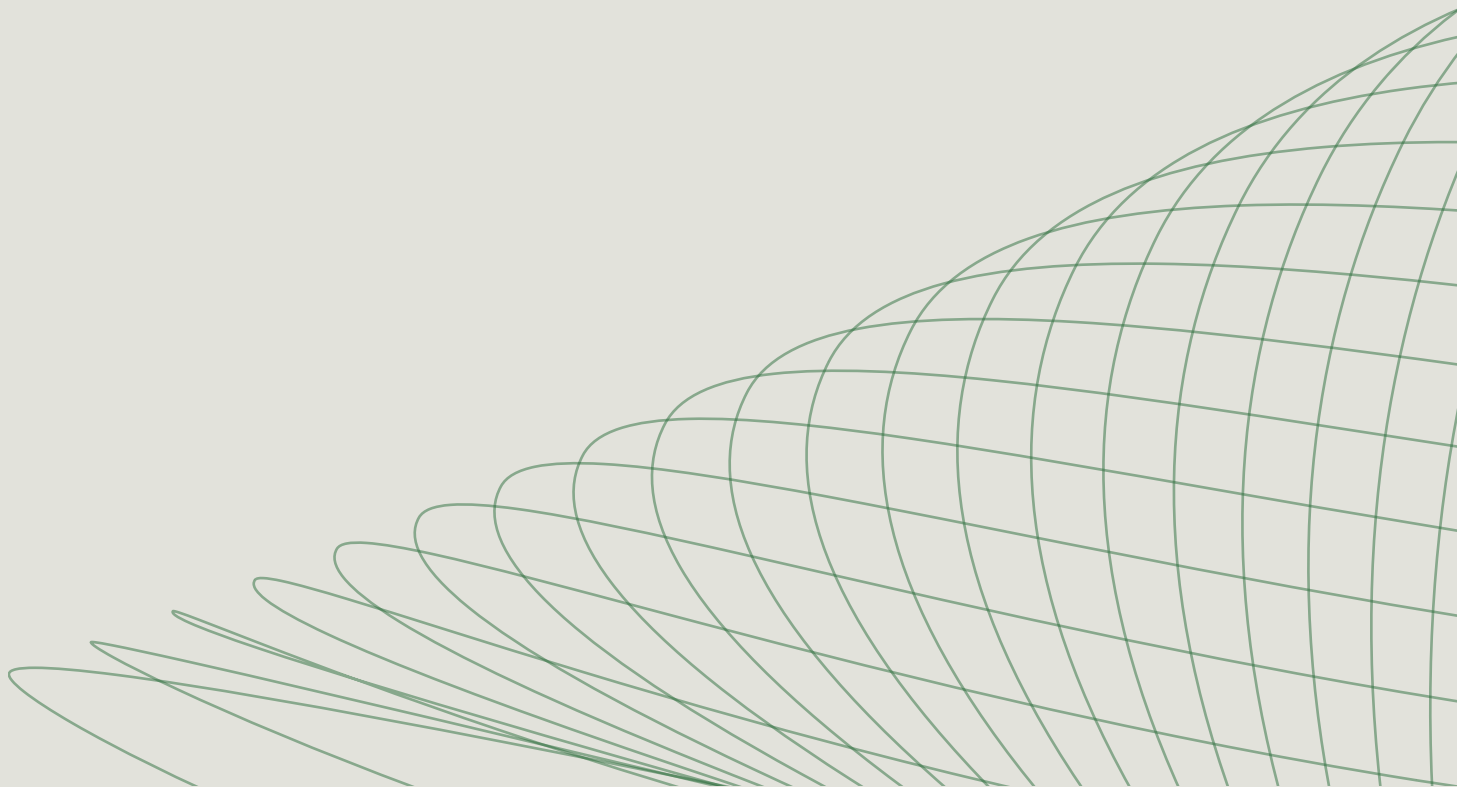




RESEARCH REPORT

State of AI Architecture 2026



Contents

Introduction	3
Insight 1: Lock-in is a fear, but not (yet) a buying criterion	4
Insight 2: Enterprises are seeking a cross-provider solution now that multi-model is the default	9
Insight 3: The agent layer is stickier than the model	11
Insight 4: Appetite for open exceeds readiness for open	14
Respondent profile	19
Geographies	19
Industries	19
Organization size	20
Respondent role	21
AI maturity	22
Conclusions	23
About Stacklok	24

Introduction

In September 2026, we surveyed 520 leaders who are responsible for their organization's use of large language models and/or AI agents. All respondents work at companies with more than 500 employees in the United States or Canada, hold a manager-or-above title, and have AI agents in pilot or production. Respondents came from twelve industries, with software (27%) and financial services (16%) the largest groups.

Most enterprise AI conversations focus on the model: which one, from whom, and at what price. This study looks at everything that gets built around the model, including the integrations, workflows, guardrails, memory and evals that turn a model into a working system, then asks how portable that system really is. We wanted to understand how enterprises are assembling their AI stacks, how much they worry about being locked in to any one model or provider, and what (if anything) would actually be hard to move.

Following is a summary of the results. As with our previous research, we are intentionally avoiding subjective commentary and sales pitches; we figure you're here because you want the benchmarking data points.

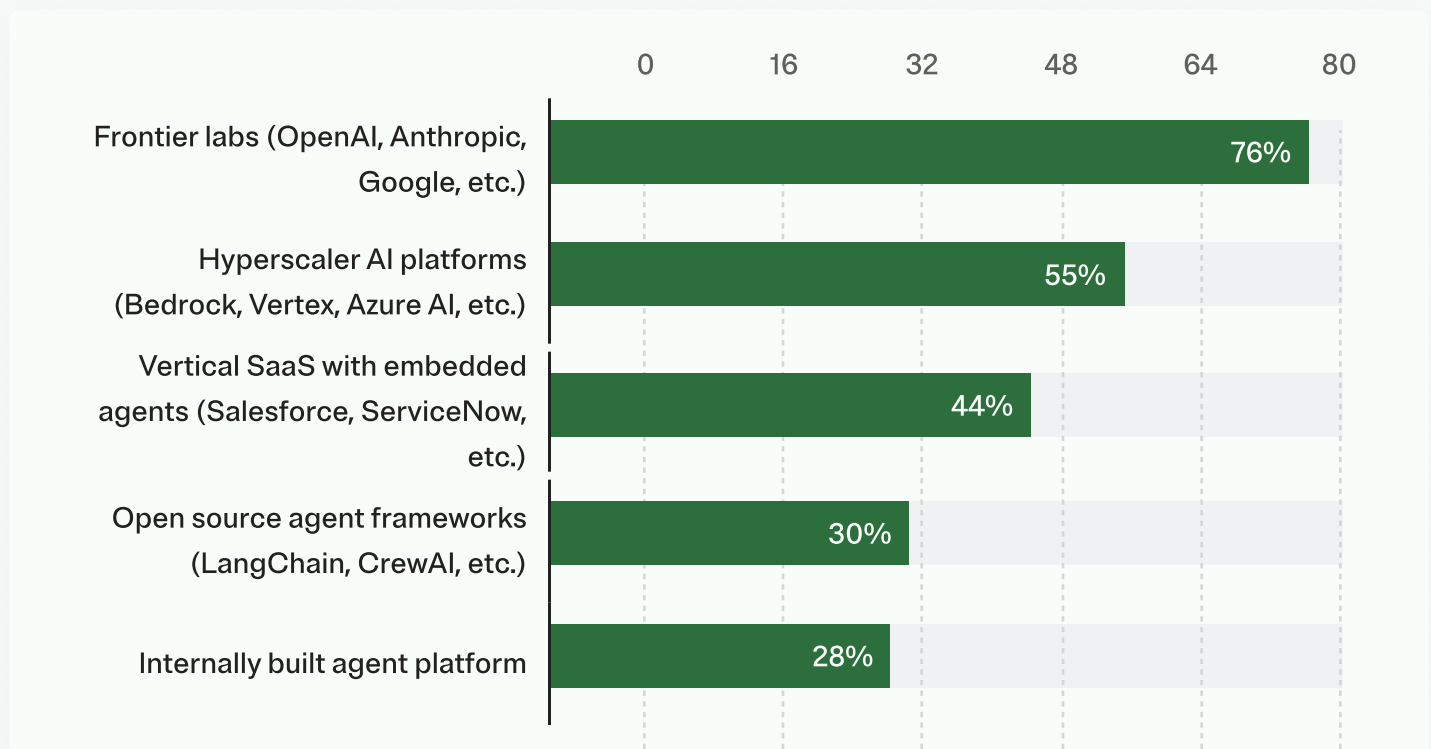
What we mean by “open architecture”

A closed, or vertically integrated, architecture is one where the layers (the model, harness, runtime, tool integrations, memory, evals and more) are supplied by a single vendor and are designed to work together on that vendor's terms. An open architecture is one where the layers around the model are built on open standards and open source software so that an enterprise can use any model from any provider, and can move between providers without having to rebuild layers.

Insight 1: Lock-in is a fear, but not (yet) a buying criterion

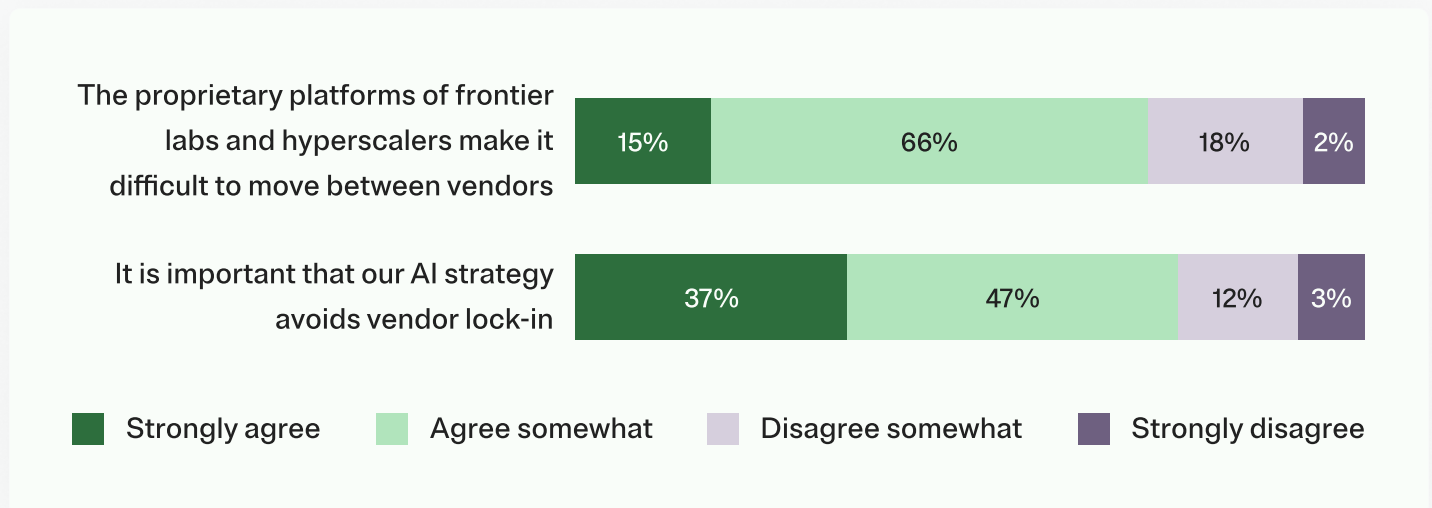
We asked respondents about the composition of their current agent stack. Frontier lab APIs are close to universal, hyperscaler platforms are in a majority of stacks, and roughly three in ten organizations have an open source agent framework (per above, an internally built agent platform). With that baseline established, we dove into the study.

“Which of the following are part of your organization’s AI stack today?”



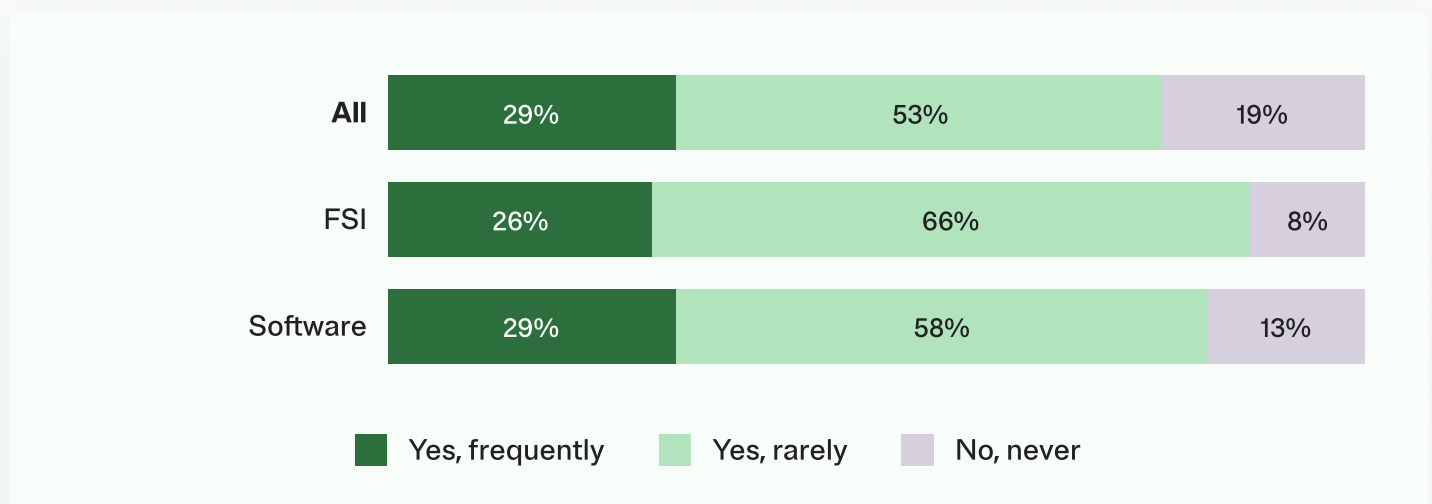
Given that a majority of respondents are utilizing vertically integrated stacks, we wanted to explore awareness and perception of lock-in risks. Respondents were nearly unanimous that lock-in exists and that it matters: 81% agree that frontier labs and hyperscalers make it difficult to move between vendors, and 84% say it is important that their AI strategy avoids vendor lock-in.

“Please indicate your agreement with each of the following statements.”



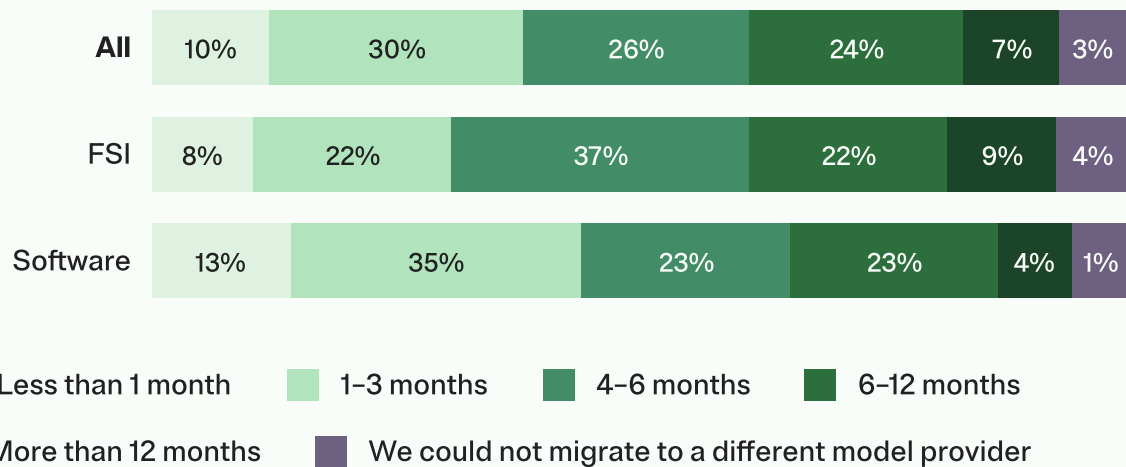
Respondents reinforced that this is not a theoretical worry. 81% report that concern about switching costs has already caused their organization to delay, scale back or restructure a migration from one model provider to another; 29% say it happens frequently. In financial services, 92% have experienced this at least once.

“Have concerns about vendor switching costs ever caused your organization to delay, scale back or restructure a migration from one model provider to another?”



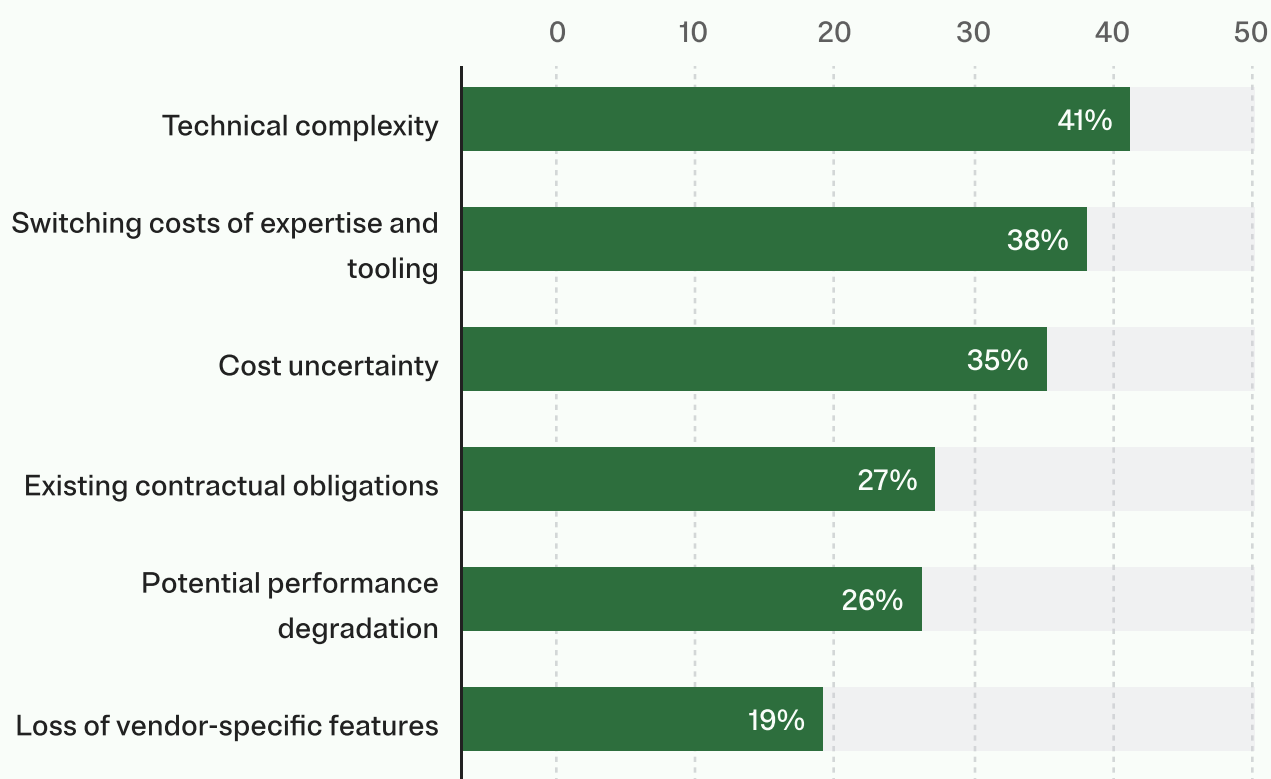
When asked to estimate how long it would take to move their current workloads to a different model provider, 57% said four months or more, and 3% said they could not migrate at all. Financial services firms expect the longest migrations: 68% estimate four months or more.

“If your team decided to move these workloads to a different model provider, how long do you estimate the migration would take?”



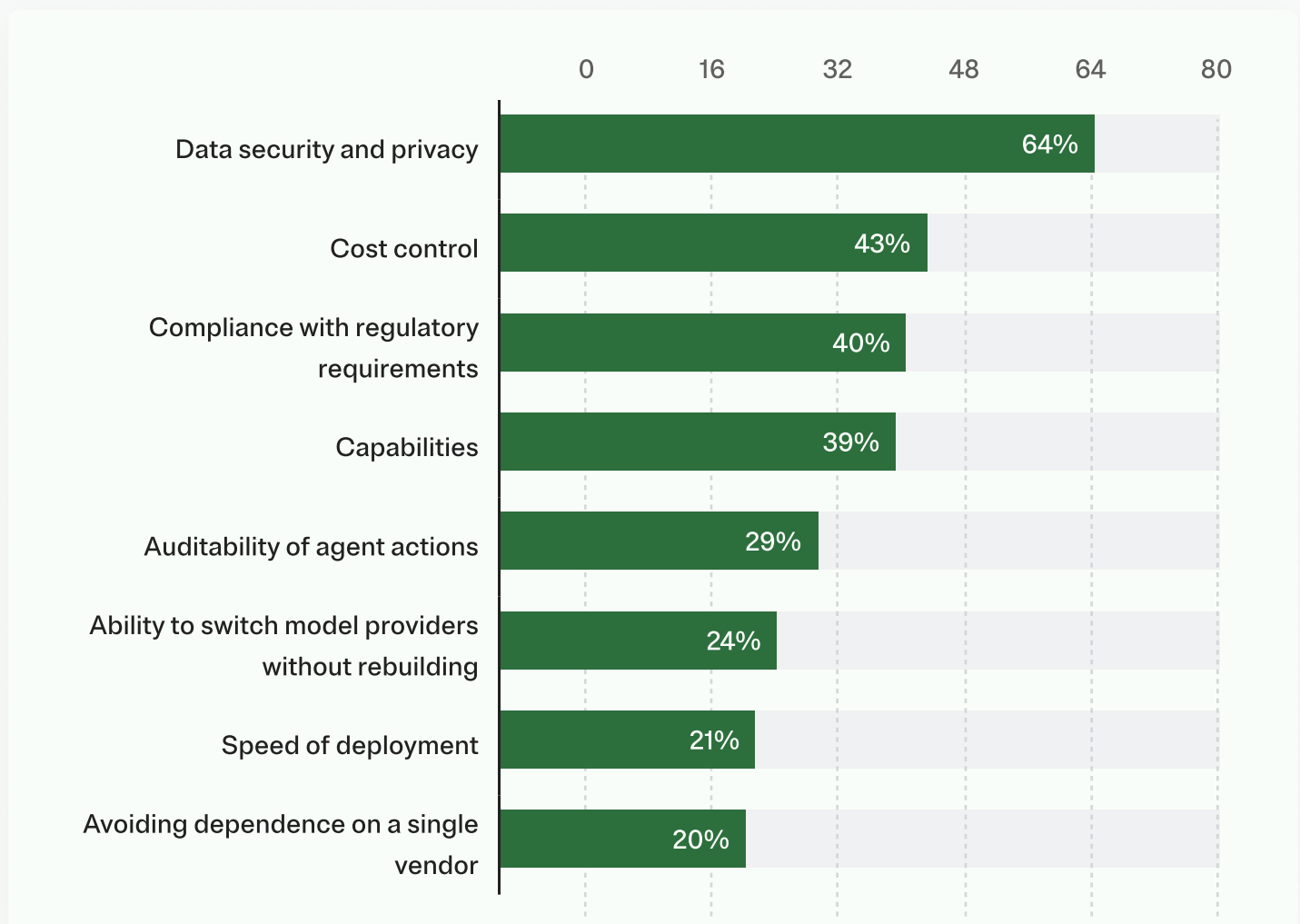
The obstacles that respondents cite are mostly engineering barriers rather than commercial ones. Technical complexity leads, followed closely by the switching cost of expertise and tooling. Contractual obligations rank fourth.

“What is the biggest barrier to switching AI model providers?
Choose up to 2 of the following.”



Here is the gap. When the same respondents were asked what matters most when selecting AI infrastructure, the ability to switch providers without rebuilding ranked sixth out of eight options, and avoiding dependence on a single vendor ranked last. Data security and privacy dominated, followed by cost control, regulatory compliance and capabilities. Lock-in is something respondents expect to suffer, but that's a problem that arrives later, whereas cost controls and compliance have to be dealt with now. This may be present bias, discounting future pain, or the tragedy of the commons, in that lock-in is likely to be someone else's problem down the line.

“What are the most important factors when your organization selects AI infrastructure? Choose up to 3 of the following.”

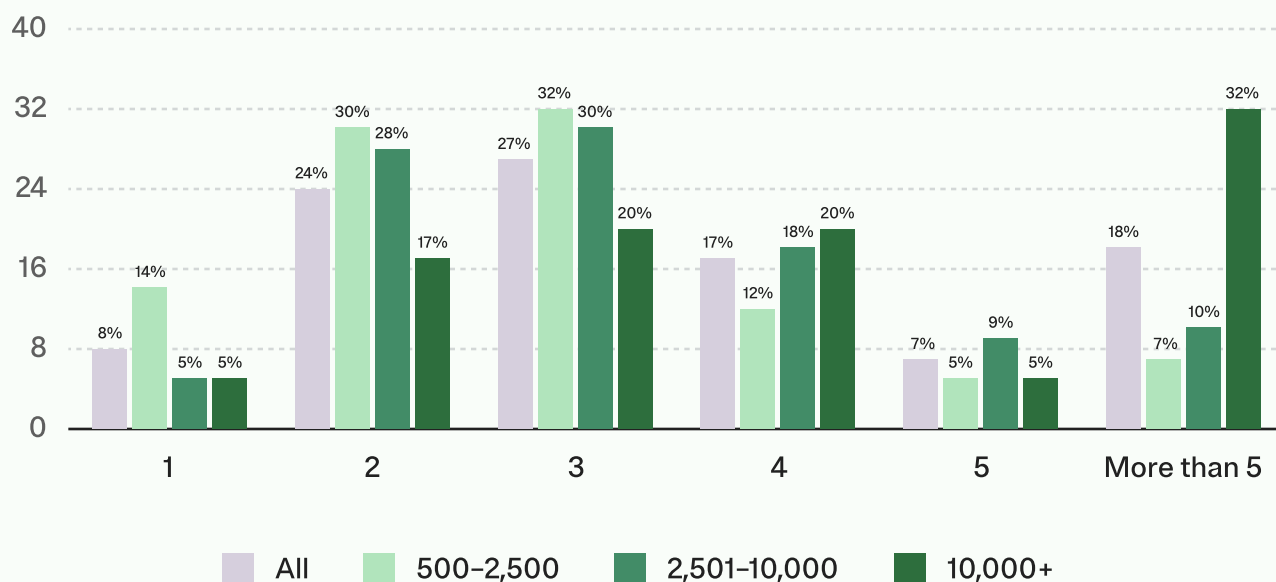


Two cuts are worth noting. Security leaders rank auditability of agent actions much higher than their peers (45% vs. 29% overall). And organizations already using three or four model providers are the most likely to rank switching ability as a top-three factor (30%), a preview of the theme in the next section. That's a nifty segue; let's go there now.

Insight 2: Enterprises are seeking a cross-provider solution now that multi-model is the default

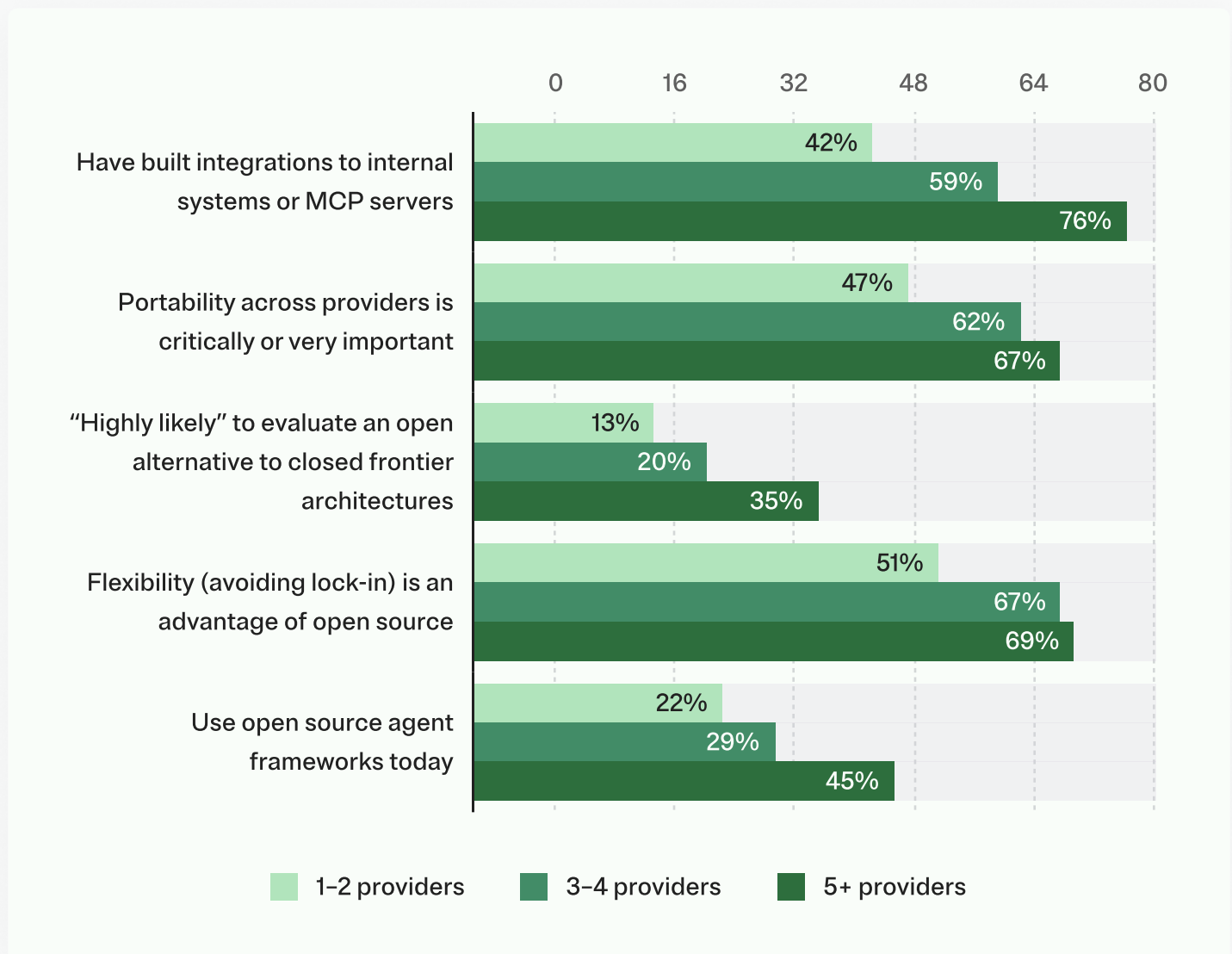
Only 8% of respondents use one model provider. Half use two or three, and nearly a quarter use five or more. The larger the enterprise, the more likely it is that they are using multiple model providers. 37% of companies with more than 10,000 employees use five or more providers, compared with 12% of companies with 500 to 2,500 employees.

“How many different model providers does your organization use in production or pilot today?”



The organizations with five or more providers are heavier users of every stack component, and they are far more likely to have built the connective layers that make a multi-model architecture work: 76% have built integrations to internal systems or MCP servers, versus 42% of organizations using one or two model providers.

“Selected results by number of model providers in use”

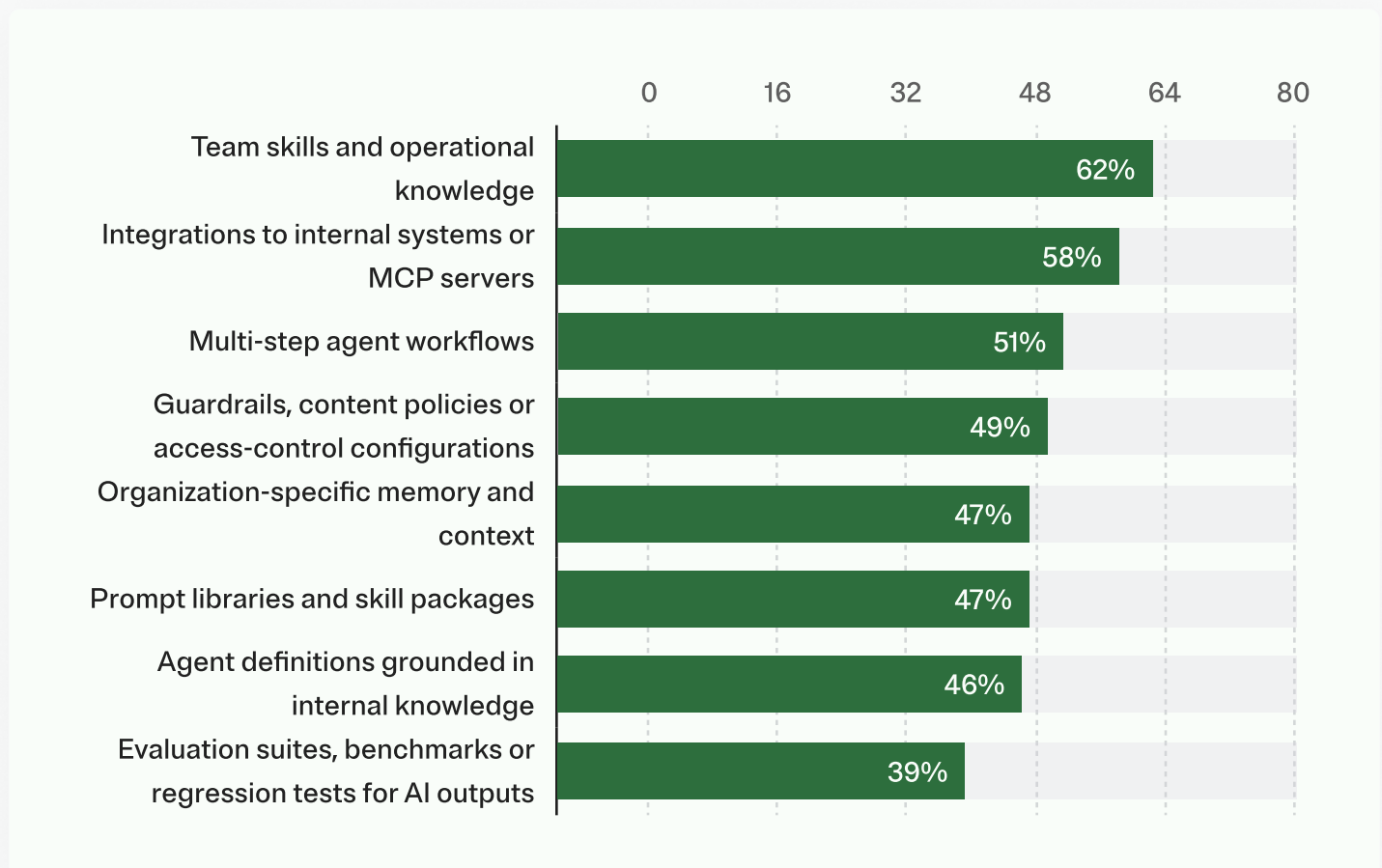


In other words, the more providers an organization runs, the more portability moves from a nice-to-have to an operating requirement; they need an architecture that treats the model as an interchangeable component.

Insight 3: The agent layer is stickier than the model

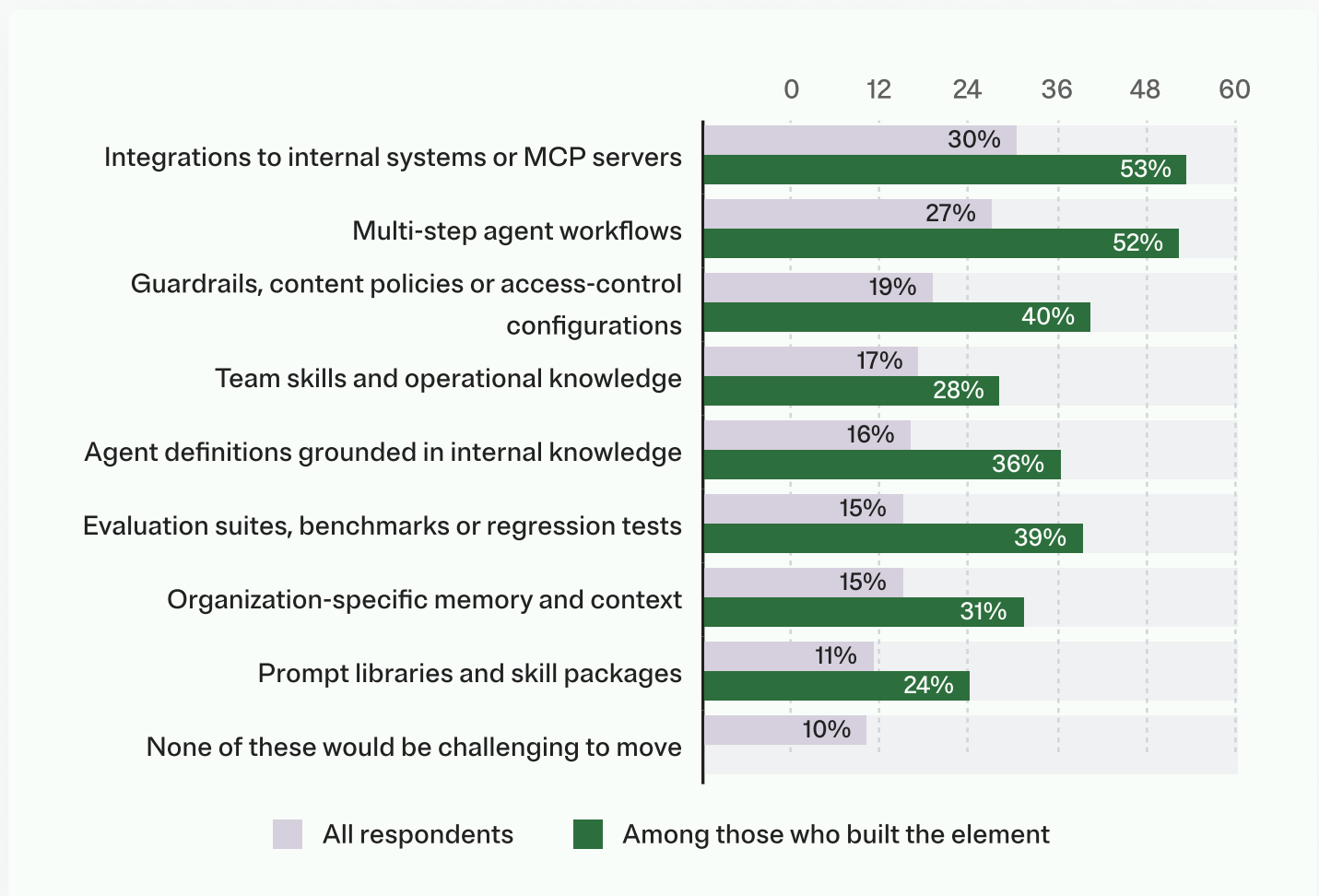
If enterprises already swap models routinely, where does lock-in actually live? We asked respondents which elements of an AI system they had built or deployed on their current models. Most have built a substantial amount around the model. At the top of the list are team skills, integrations to internal systems or MCP servers, and multi-step agent workflows.

“Which of the following elements of an AI system has your organization built or deployed on your current models?”



We then asked which of those elements would be most challenging to move to another model provider. Integrations to internal systems and MCP servers ranked first, followed by multi-step agent workflows. Prompt libraries, the element most closely tied to a specific model's behavior, ranked last. Only 10% of respondents said none of these elements would be hard to move.

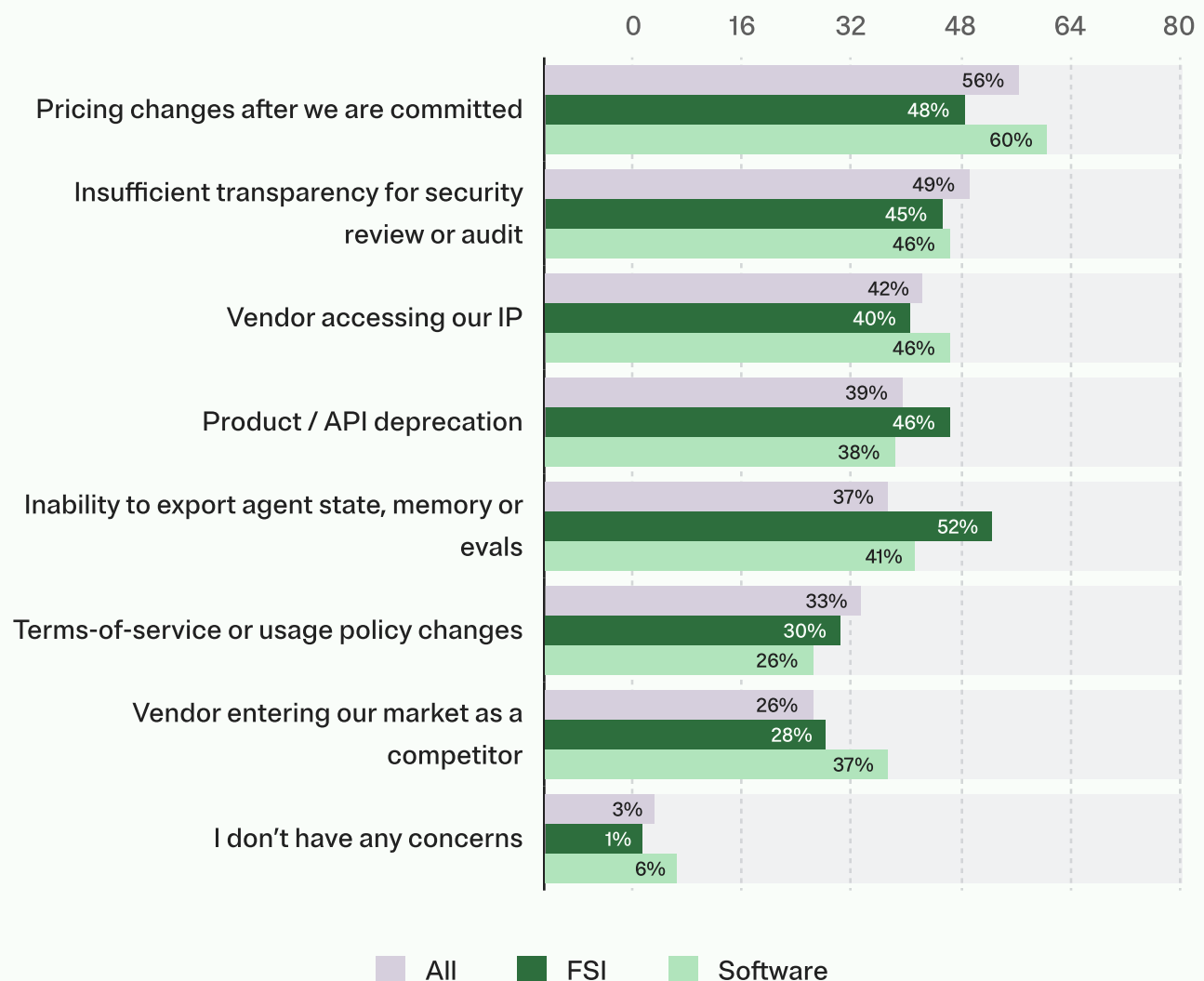
“Which part of your existing AI system would be most challenging to move to another model provider? Choose up to 2 of the following.”



Among organizations that have built integrations or multi-step workflows, more than half name that element as one of their two hardest things to move. The work that connects an agent to the enterprise is the work that is hardest to take with you.

This shows up again in what respondents worry about when they build on a frontier lab or hyperscaler platform. Pricing changes after commitment is the most common concern overall, but 37% worry specifically about the inability to export agent state, memory or evaluations, and 39% fret over product or API deprecation. Only 3% have no concerns.

“What concerns do you have about building on frontier lab or hyperscaler platforms?”

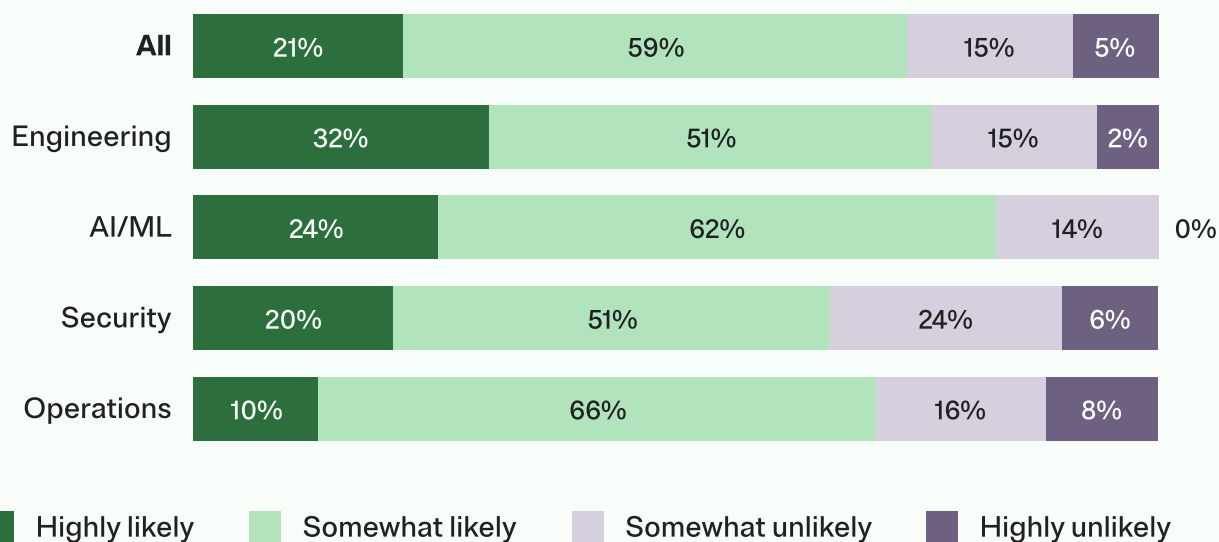


In financial services, the inability to export agent state, memory or evals is the single most common concern; whereas retailers are worried about pricing and vendors entering their market.

Insight 4: Appetite for open exceeds readiness for open

Given all of the above, it is not surprising that interest in an open alternative is broad. 80% of respondents say they would be likely to evaluate an open source alternative to the vertically integrated, closed architectures offered by frontier labs, and one in five is highly likely. Engineering leaders are the most enthusiastic (32% highly likely); operations leaders the least (10%).

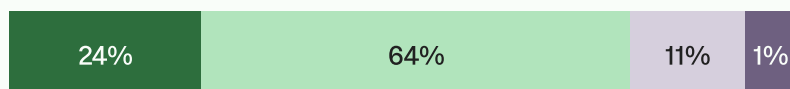
“If there was an open source alternative to the vertically-integrated and closed architectures offered by frontier labs, how likely would your organization be to evaluate it?”



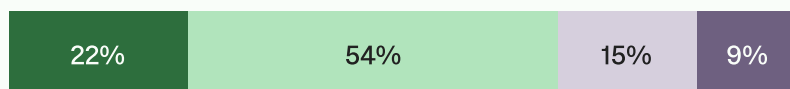
The appetite extends to control more generally. 88% agree that their team expects to take more control of its AI architecture as it gains experience, rather than continuing to rely so heavily on current vendors. Among AI/ML leaders that figure is 97%. And 76% describe themselves as very interested in open source alternatives for AI architectures.

“Please indicate your agreement with each of the following statements.”

Our team expects to take more control of our AI architecture as we gain experience rather than relying so heavily on current AI vendors



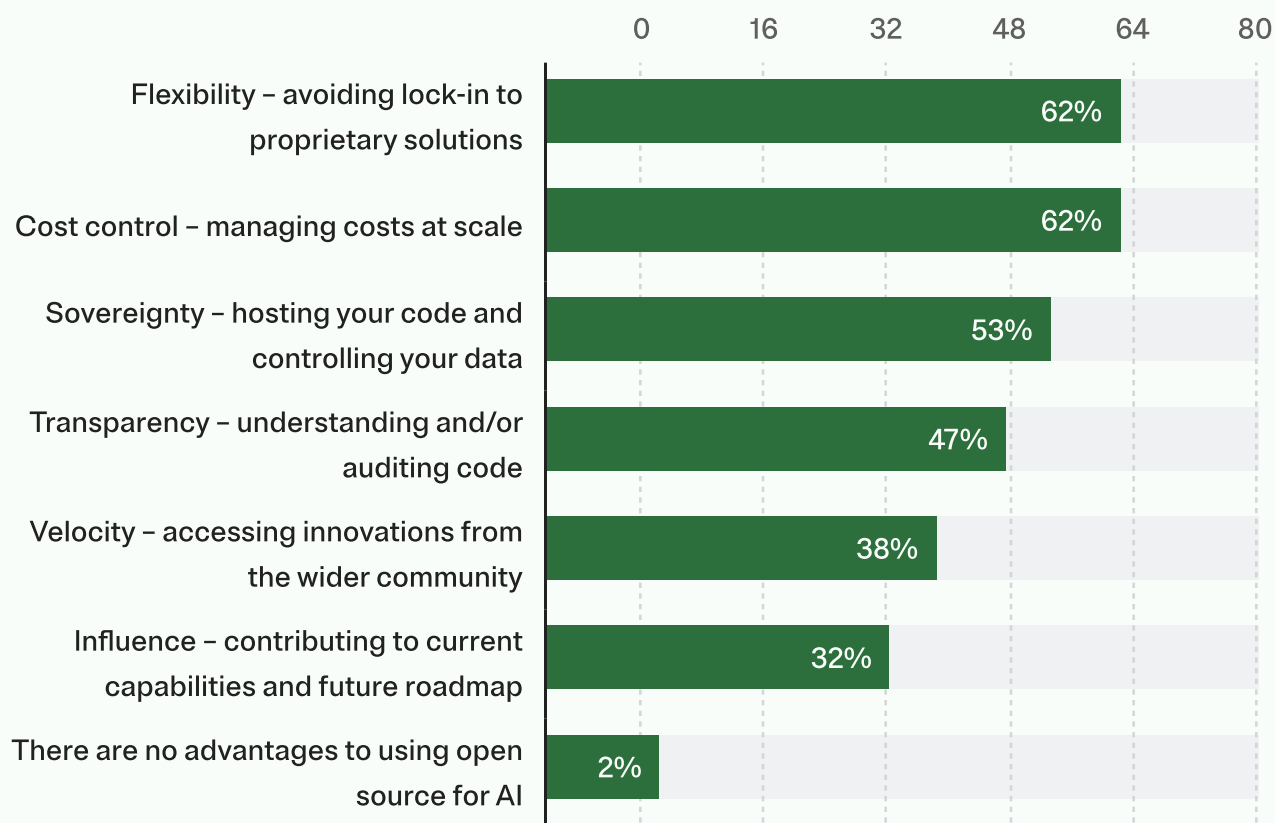
I am very interested in open source alternatives for AI architectures



Strongly agree Agree somewhat Disagree somewhat Strongly disagree

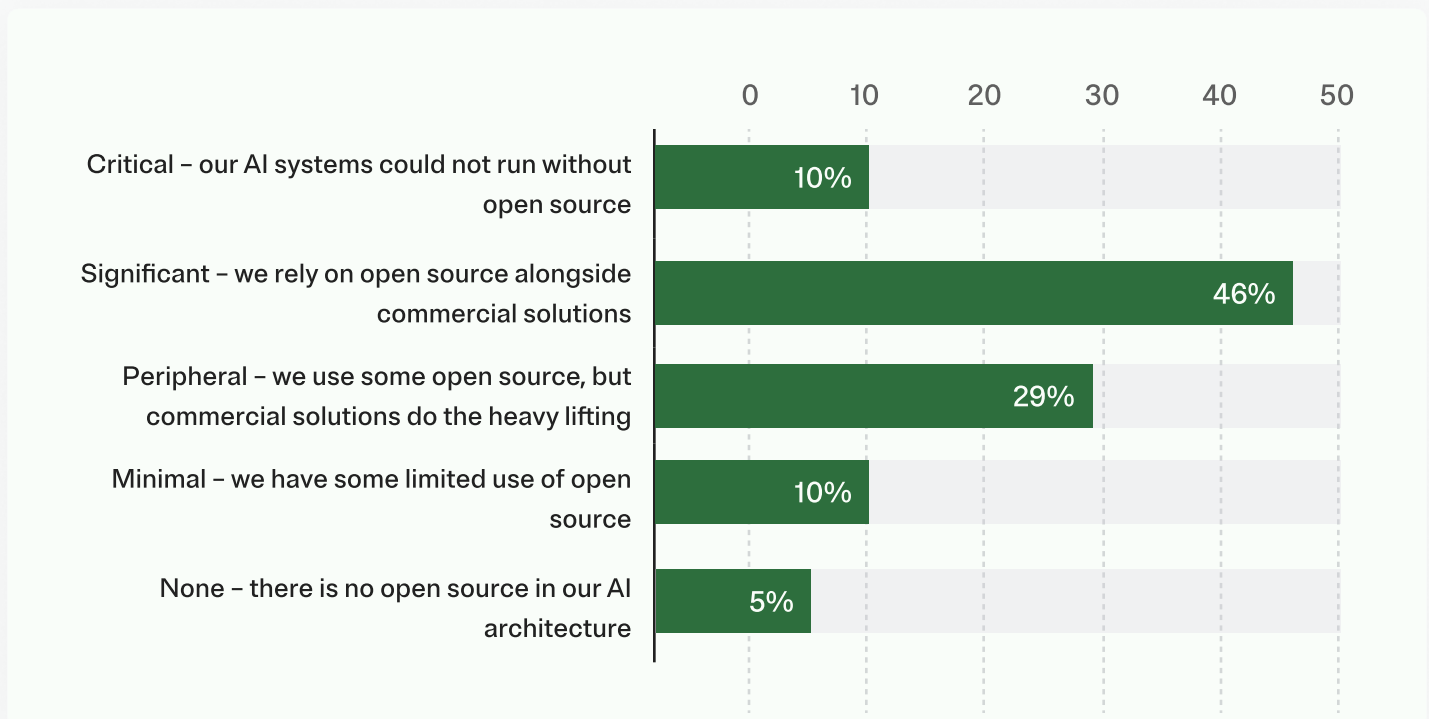
Respondents are clear about their motivations. Flexibility and cost control tie as the leading advantages of a more open architecture, with sovereignty over code and data close behind.

“In your opinion, what advantages can be gained from the use of open source in your AI architecture?”



However, respondent practices lag some of their intentions. Today, open source is critical or significant in the AI architecture of 56% of respondents; for the other 44%, commercial solutions do the heavy lifting.

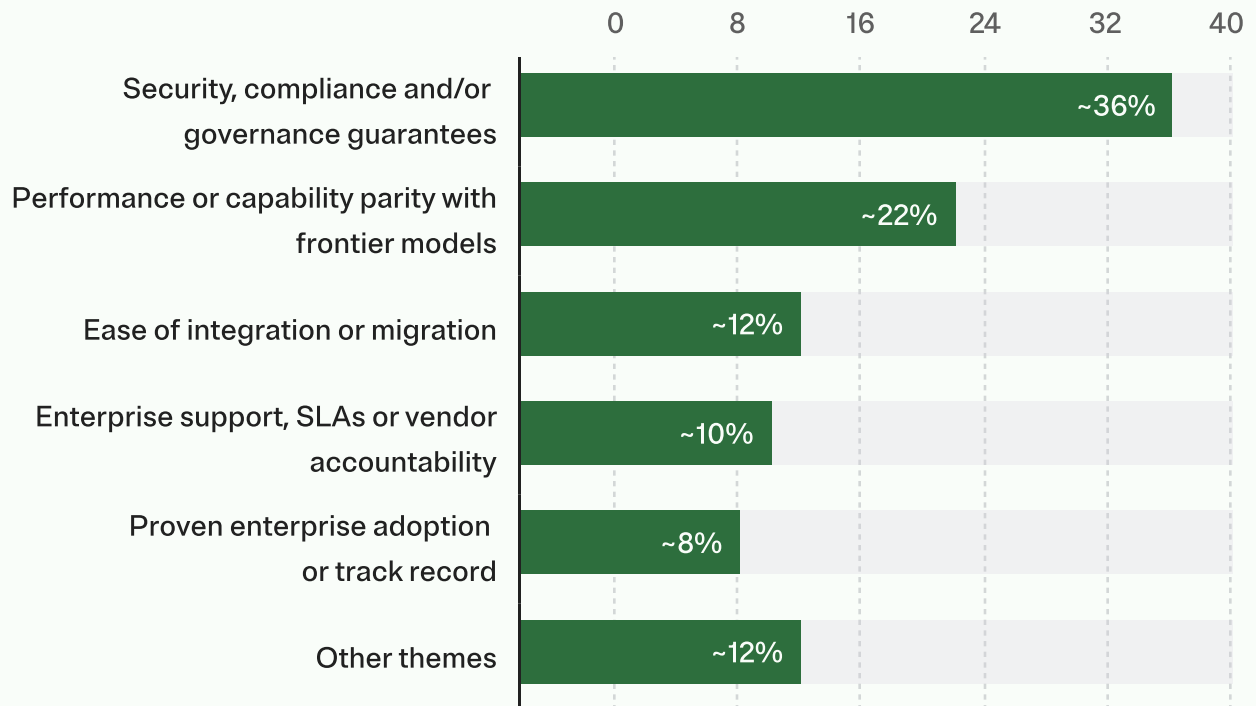
“Which of the following best describes the role of open source software in your AI architecture today?”



We asked every respondent to complete the sentence “We would move more AI workloads to an open architecture if...” and we received 277 substantive answers. Coded by theme, roughly four in ten mention security, compliance, privacy or governance. The next most frequent theme was performance or capability parity with frontier models, followed by ease of integration or migration.

Theme in “We would move more workloads to an open architecture if it offered ...”

% of responses



Respondent profile

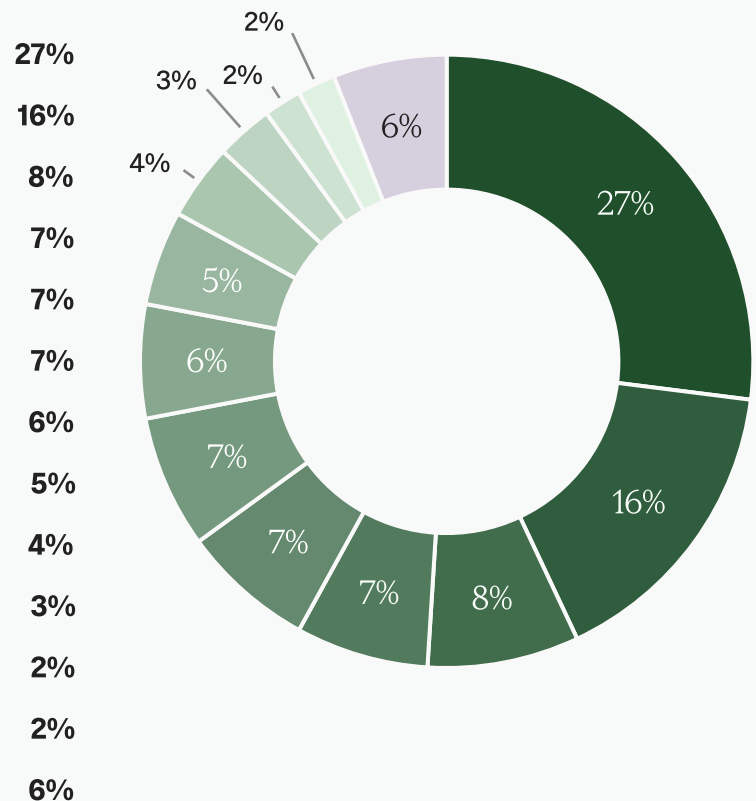
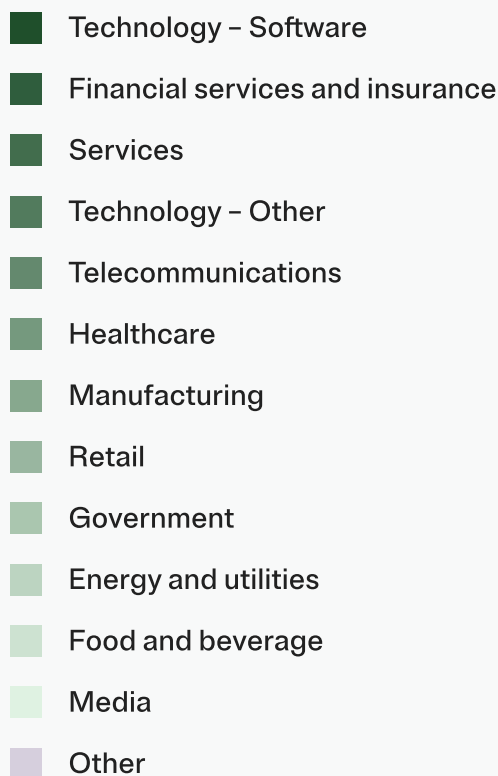
Geographies

This was a North American study. All respondents were based in the United States or Canada.

Industries

The study did not exclude any industries. Software and financial services together account for 43% of the sample and are the two industries we occasionally drill into in the report (given sufficient sample).

Primary industry

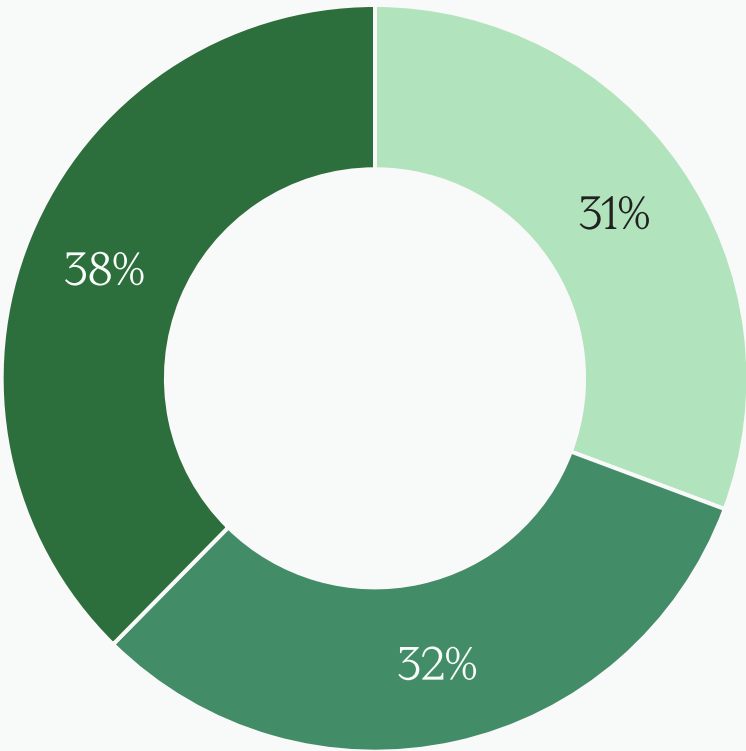


Organization size

The study targeted mid-sized and large enterprises. Seventy percent of respondents work at companies with more than 2,500 employees, and 38% at companies with more than 10,000 employees.

Employees

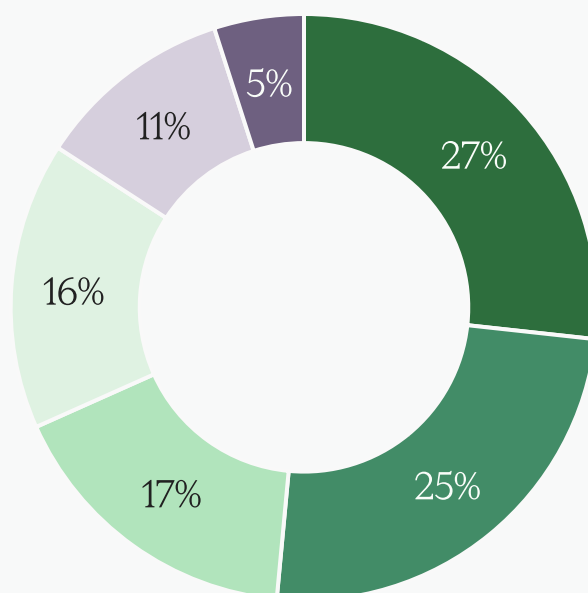
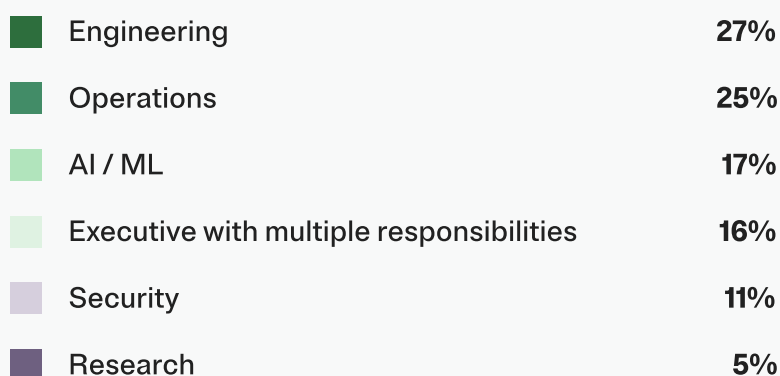
500 – 2,500 employees	31%
2,501 – 10,000 employees	32%
More than 10,000 employees	38%



Respondent role

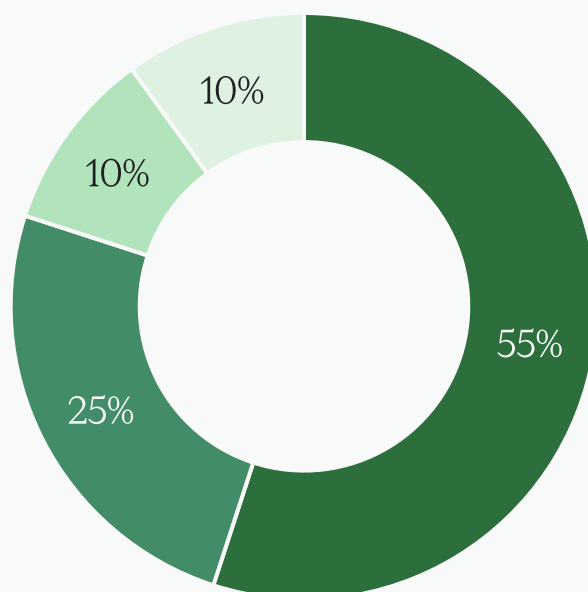
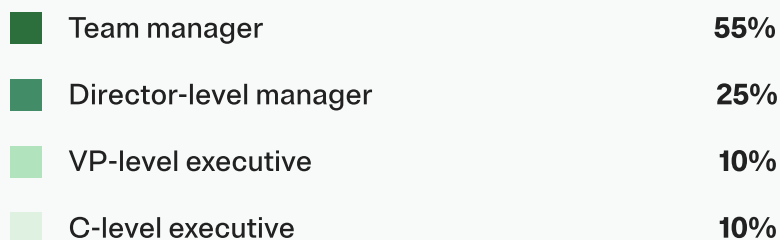
Respondents were screened for direct responsibility over their organization's use of LLMs or AI agents. They sit across the functions that typically share ownership of an AI platform, with engineering and operations the largest groups.

“What function best describes your role?”



Forty-five percent of respondents hold director-or-above titles.

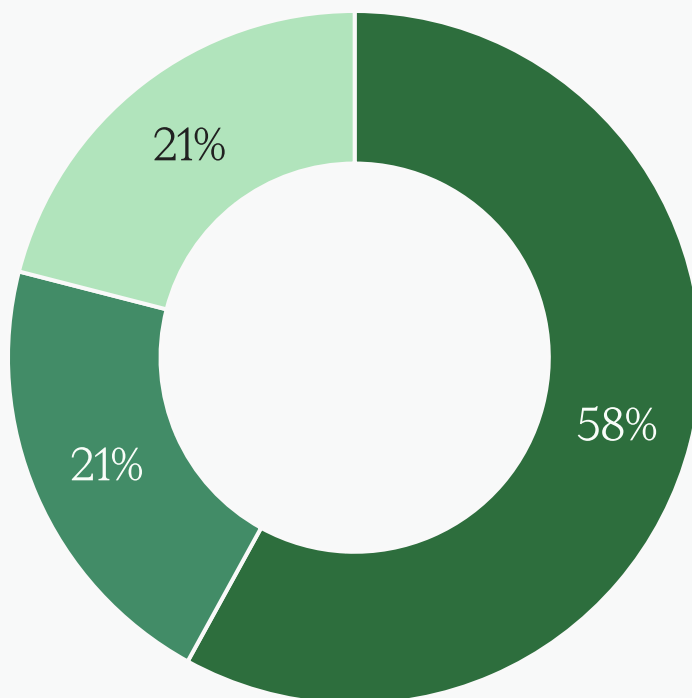
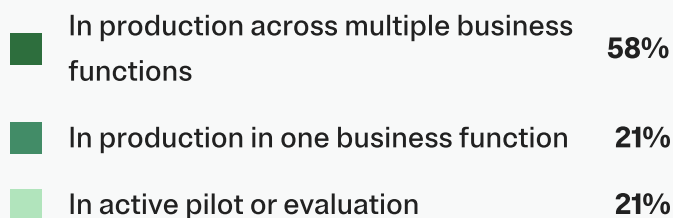
“Which of the following best describes your current job title?”



AI maturity

Anyone whose organization was still in early exploration, or had no AI activity, was filtered out of the study. The remaining respondents are well past experimentation: 58% run LLMs or AI agents in production across multiple business functions.

“Which best describes your organization’s current use of LLMs or AI agents for building or delivering technology solutions for employees, customers or partners?”



Conclusions

Enterprise AI leaders are already multi-model: 92% use more than one provider and a quarter use five or more. They overwhelmingly recognize that proprietary platforms make systems hard to move, and most have already delayed or restructured a migration because of it. Yet portability and vendor independence still rank near the bottom of their stated selection criteria, behind security, cost, compliance and capabilities.

The data suggests why. Lock-in is not accumulating at the model layer, which respondents swap with some regularity, but in the agent layer around it: the integrations to internal systems and MCP servers, the multi-step workflows, the guardrails and the organization-specific memory that turn a model into a working system. Those are the elements enterprises have invested in most heavily and expect to find hardest to move. For example, in financial services, the inability to export state is now the leading concern about building on a frontier lab platform.

The appetite for an alternative is broad. Four in five would evaluate an open architecture, and nearly nine in ten expect to take more control of their AI architecture over time. What stands between interest and adoption is assurance; respondents expect an open architecture to meet exacting standards for security, compliance, support and performance, and that lets them keep using the frontier models they already rely on. Where those conditions are met, the organizations furthest along in their AI journey are the ones most ready to move.

About Stacklok

Stacklok offers an open architecture as an alternative to the vertically integrated stacks from frontier labs. Our customers own their harness, context, tools, policies and more, so they can move between models and providers to manage costs and optimize performance. Our platform governs the two doors in-and-out of enterprise AI architecture with an MCP gateway and LLM gateway. We offer enterprises a consistent approach to identity, auth, observability and audit across their agent stack, and a way to scale to all categories of knowledge workers using their existing, trusted Kubernetes infrastructure. Stacklok is founded and led by the co-creators of Kubernetes, maintains the official MCP registry, and builds with the community.

Learn more about Stacklok:

www.stacklok.com

Explore our GitHub repo:

<https://github.com/stacklok>

Engage with us via Discord:

<https://discord.gg/stacklok>