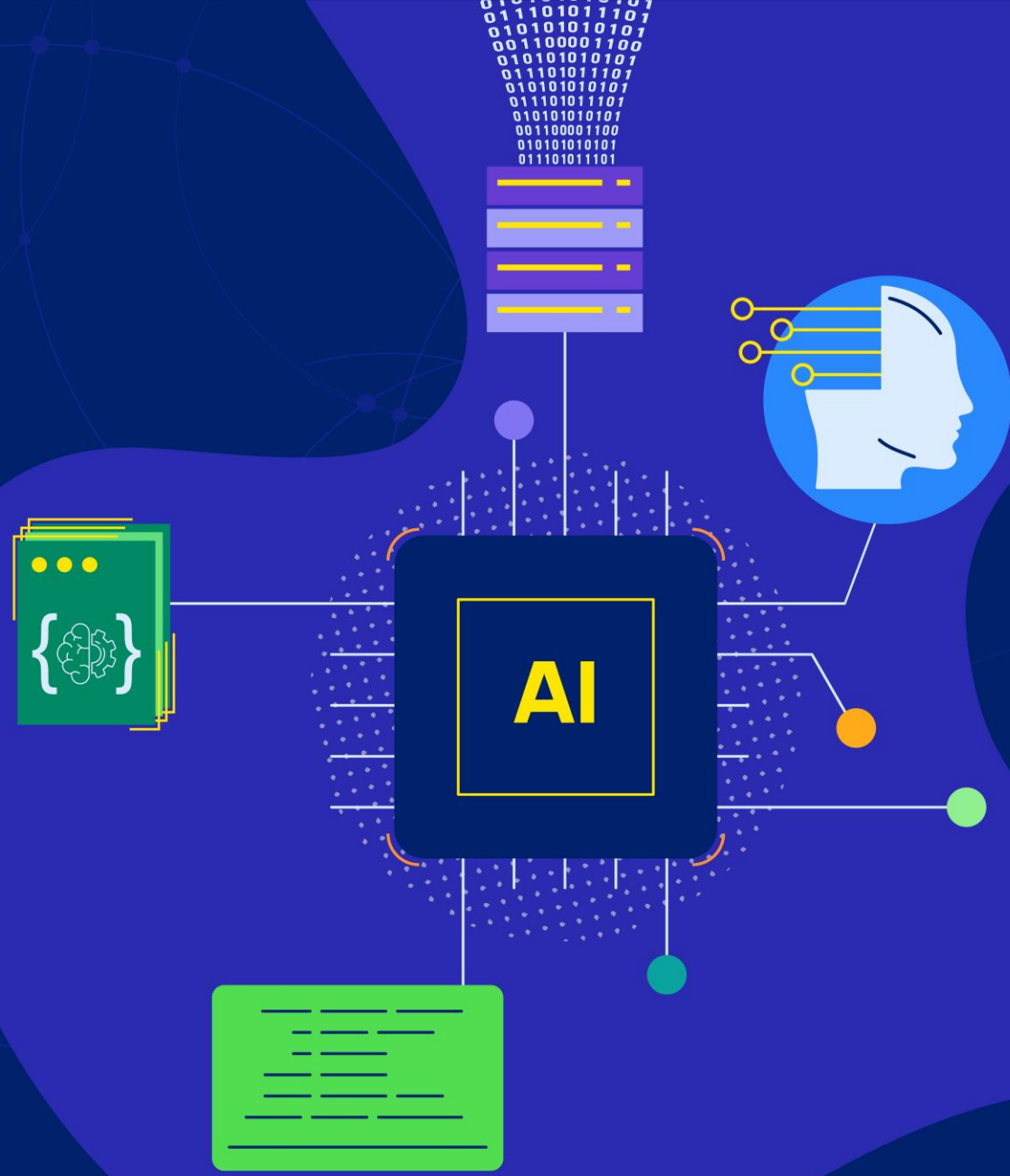
November 2026

State of AI Engineering in 2026:

From AI Adoption to Engineering Maturity.



Progress® | /DATA

ABOUT THE AUTHORS

myriam.boualami@slashdata.co



Myriam Boualami

Market Research Analyst

Myriam investigates how new technologies shape real-world engineering practices and the future of software development. Her background includes over four years as a research engineer in a computational social science lab, where she focused on information and communication technologies.

bleona.bicaj@slashdata.co



Bleona Bicaj

Principal Research Consultant

Bleona is a research consultant, enthusiastic about product strategy and behavioural science. She holds a Master's in Economic and Consumer Psychology. With more than 6 years of professional experience as an analyst, she has worked across quantitative and qualitative research studies, turning complex data into clear narratives that inform better products, smarter investments, and long-term growth.

Progress | /DATA

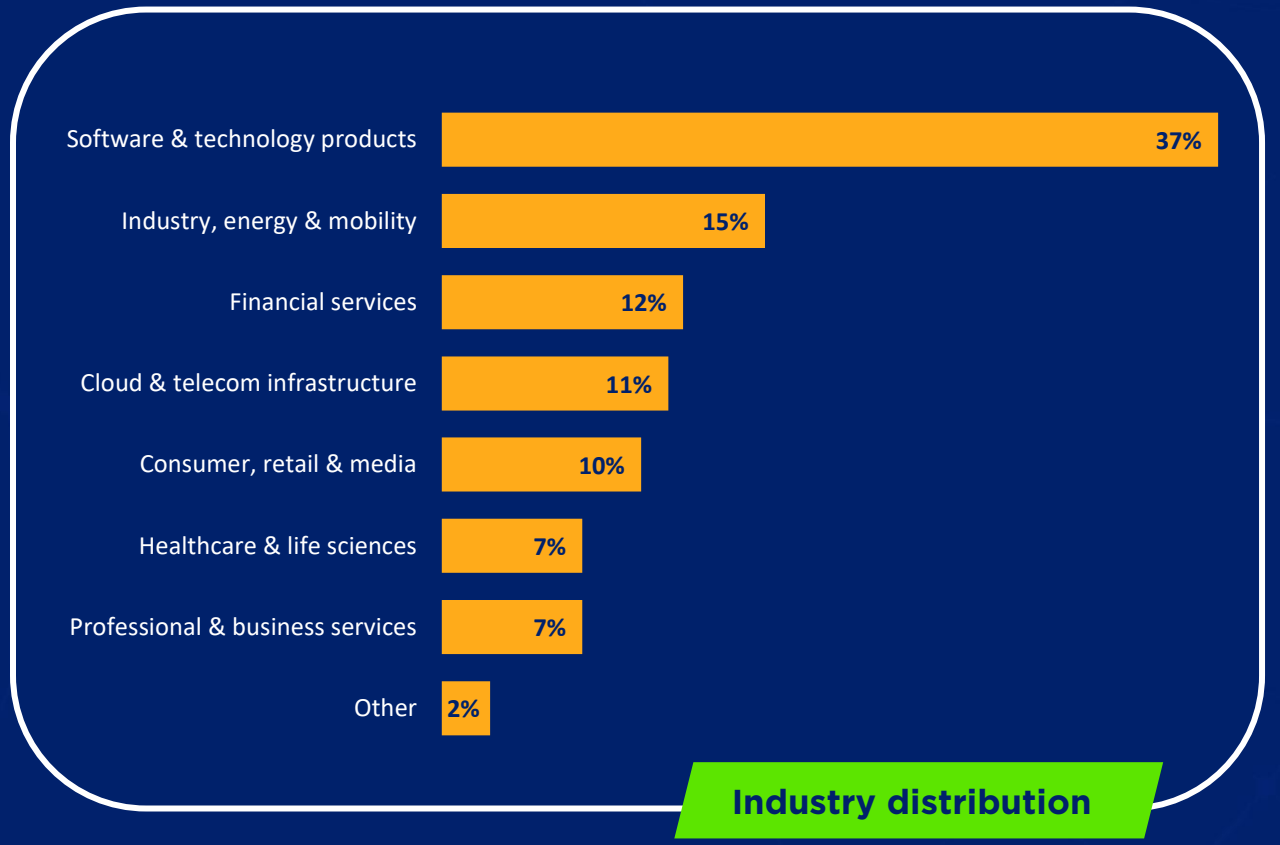
TABLE OF CONTENTS

Introduction	4
Report highlights	8
The state of AI engineering	10
Profiling mature organisations	26
Keeping production AI under control	40
The future of software development	53
Conclusions and recommendations	62
Methodology	66

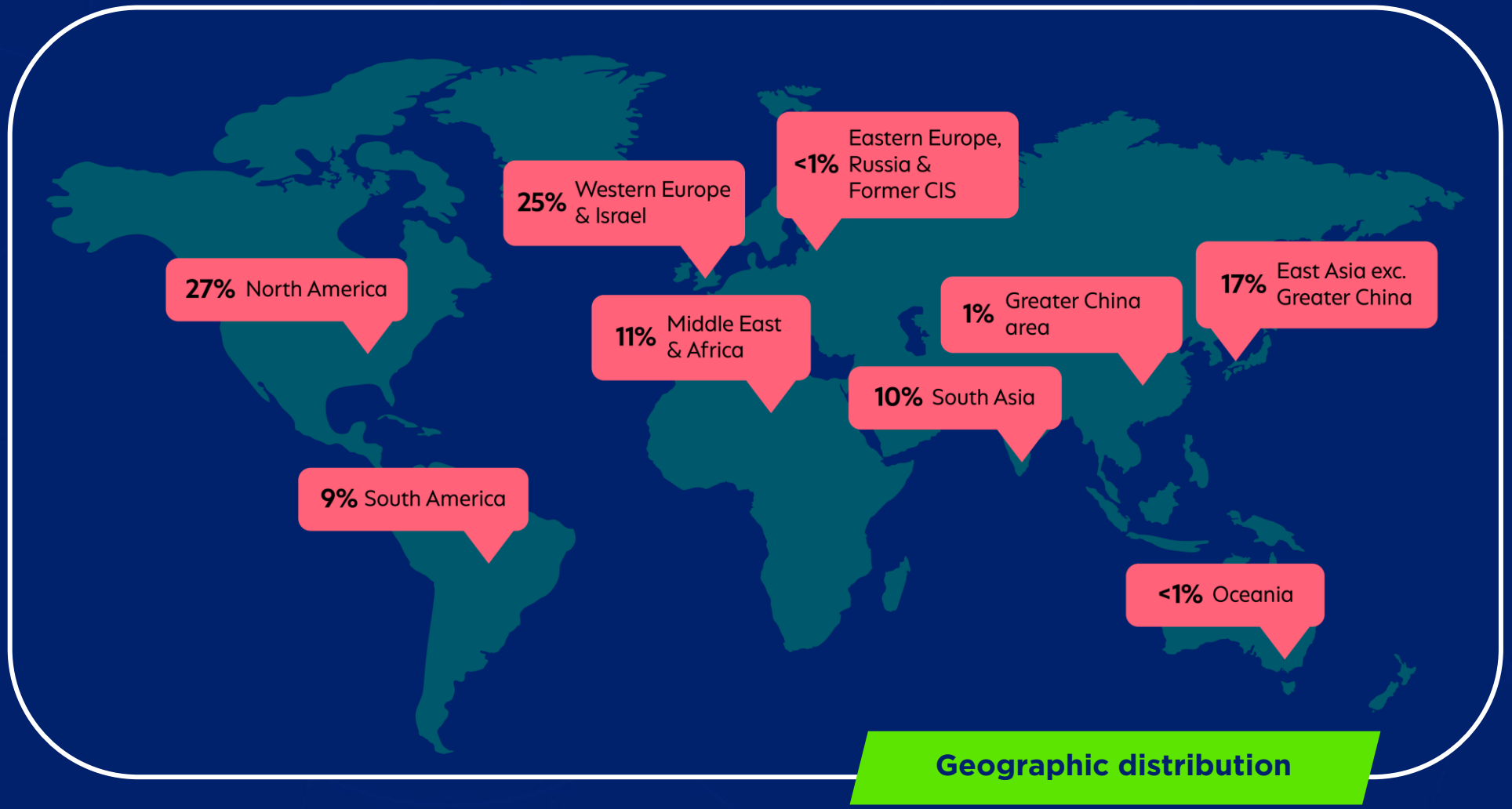
Progress® | /DATA

Who we surveyed:

- **862** qualified software professionals with 3+ years professional software-development experience required.
- **100%** of respondents directly involved in AI engineering beyond everyday use of AI coding tools.
- **143** technology leaders (CTOs, VPs and Directors).
- **719** engineers / other technology practitioners.
- **Q3 2026** global survey conducted by SlashData for Progress Software.



Industry distribution



Geographic distribution

ABOUT THIS REPORT

Software organisations are increasingly integrating LLM-based features into applications and development workflows. We refer to those practices as AI engineering. Fewer can say how far AI engineering work has progressed or what separates the organisations that run AI in production at scale from those still running pilots. This report addresses that gap.

Throughout, an AI capability means a feature, function, system, or workflow that uses a large language or foundation model to perform or support a specific task. It may be developed internally or rely on third-party models and services. Conventional AI adoption metrics cannot distinguish an organisation running isolated experiments from one operating AI capabilities reliably at scale. To measure those differences, we built the AI Engineering Maturity Index. The index scores each organisation from 0 to 1 across five dimensions. **The 0–1 scale reflects how informal a practice is (0) versus how systematically scalable it is (1), which we call the maturity scale.** We measure a total of five aspects of AI engineering maturity:

1. **Adoption and Integration:** how far AI capabilities have progressed from research to scalable production and how widely AI supports the development lifecycle.
2. **Governance and Control:** how models are sourced, which safeguards apply in production and to AI agents, and how AI costs are tracked.
3. **Evaluation and Monitoring:** how rigorously organisations evaluate AI capabilities before release and monitor them afterwards and what extra checks apply to AI-generated code.
4. **People and Enablement:** how organisations support and guide engineers in building AI skills, build those skills into roles, and assess whether AI tools improve engineering work.
5. **Data and Architecture:** how organisations manage the models they use, prepare internal data for AI, and adapt their architecture to run AI capabilities.

The overall index is the mean of the five dimension scores. It places each organisation in one of five equal-width stages: explorers (below 0.2), experimenters (0.2 to 0.4), builders (0.4 to 0.6), operators (0.6 to 0.8), and AI-native (0.8 and above).

The report follows four chapters. Chapter one establishes where organisations stand. Chapter two asks which organisations and practices are associated with higher maturity. Chapter three examines whether evaluation, monitoring, and controls keep pace with production AI and agent autonomy. Chapter four looks at what organisations report getting back from AI and what developers expect their work to become.

REPORT HIGHLIGHTS

- AI engineering is concentrated in the builder stage, with systematic production practices less widespread. 56% of organisations sit in the builder stage of the AI Engineering Maturity Index, 25% in the experimenters stage, and 16% in the operators stage. Fewer than 1% reach the AI-native stage. →

- AI adoption is more advanced than the governance practices needed to manage it. Adoption and integration is the most advanced dimension of the index, with an average score of 0.57. Governance and control is the least advanced, at 0.39. This gap highlights an imbalance between implementing AI capabilities and establishing systematic oversight and operational controls. →

- Shared AI architecture is strongly associated with broader AI engineering maturity. Organisations with an AI gateway or model abstraction layer score 33% higher on the index than those without one. →

- Several production safeguards remain relatively uncommon despite their strong association with AI engineering maturity. Six of the eight safeguards tracked combine an above-average association with maturity and below-average adoption. →

- Organisations permitting greater agent autonomy generally report more controls, but unintended agent actions remain a concern. The share of organisations reporting that an AI agent took an unintended action rises from 7% where agents only suggest or recommend to 25% where agents plan and execute objectives. →

REPORT HIGHLIGHTS

- More mature organisations report a broader range of AI-related incidents. Operators report an average of 2.4 incident types in the past 12 months, against 1.2 for experimenters. The most common are AI costs exceeding expectations (34%) and AI-generated code causing a security or quality issue (33%). →

- Positive perceptions of AI's return are more common than negative ones. 42% of respondents say benefits exceed costs, against 19% who say costs exceed benefits. Only 13% cite proving the return on investment as a top barrier to scaling, against 43% for security and privacy. →

- Writing code by hand is the only developer task respondents agree will take less time over the next three years: 64%. Control over technical debt is the outcome least often reported as improved (35%) and the least associated with maturity. →

- Fine-tuned models are more prevalent in customer-facing software than in internal applications or business workflows. 46% of developers in organisations building AI into customer software use vendor-hosted, fine-tuned models, compared with 30% for internal software. AI in business processes and workflows relies most on vendor-hosted, off-the-shelf models (63%). →

- Larger organisations are considerably more likely to demonstrate advanced AI engineering maturity. 28% of developers in organisations with 1,001 to 10,000 employees are operators, compared with 4% in those with 21 to 100 employees. Organisation age shows no comparable gradient. →

THE STATE OF AI ENGINEERING

01

Progress® | /DATA

Maturity across the index

Organisations have made AI adoption more systematic than the governance around it: deployment leans repeatable, governance leans ad hoc. Indeed, adoption is the most advanced dimension of AI engineering, and governance the least. Across all 862 respondents, Adoption and Integration averages at 0.57. Evaluation and Monitoring follow at 0.50, People and Enablement at 0.48, and Data and Architecture at 0.43. Governance and Control scores 0.39, 0.18 below adoption.

The ordering suggests organisations put more effort into pushing AI into production than into building the controls around it. That sequence is common with new technology. Teams prove value first and formalise later. This carries a cost because the governance dimension covers what happens when an AI capability in production misbehaves: whether it can be switched off, whether its outputs can be traced, and who owns it. We return to this later in the report.

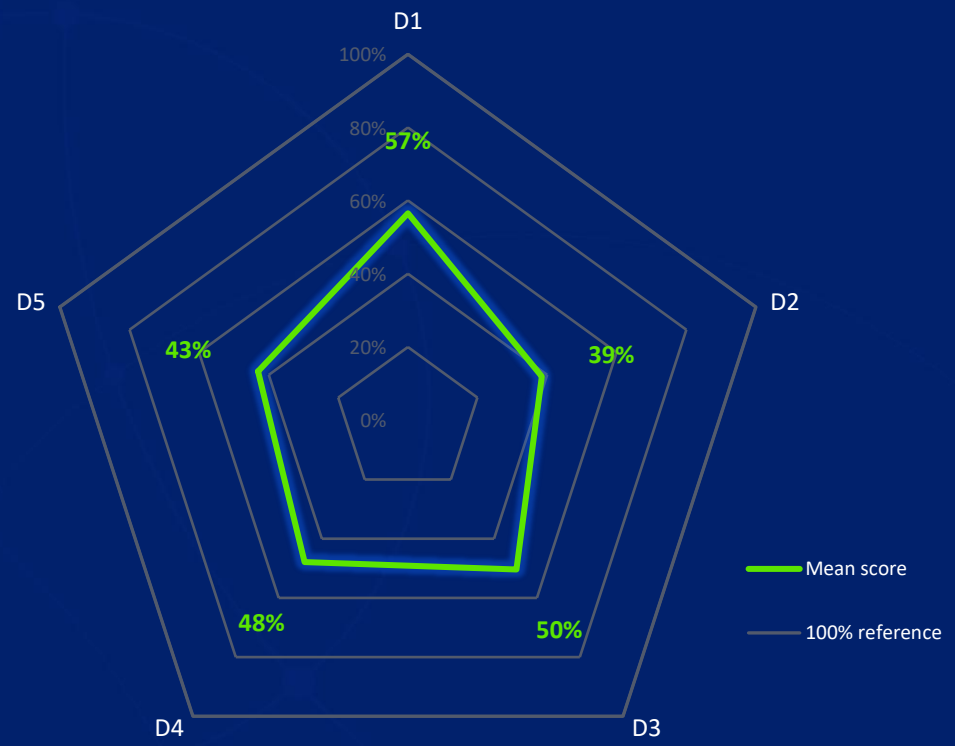
Most respondents sit in the middle of the index. 56% are builders, defined by repeatable practices for building and deploying AI. A quarter (25%) are experimenters, defined as building fragmented AI pilots, and 16% are operators, defined as systematic production engineering for AI. Explorers account for 2%, defined as individual or experimental AI activity, and AI-native respondents account for fewer than 1%, defined as established, organisation-wide AI engineering.

By our maturity scale, most respondents are concentrated in the middle, with more respondents on the less mature side than the more mature side. AI engineering has moved beyond fragmented pilots, but most organisations have not yet reached systematic production engineering. This report is accompanied by an online interactive AI Engineering Maturity Self-Assessment, built directly from the survey instrument, and readers are welcomed to self-assess their maturity here.

Adoption and integration is the most advanced dimension of AI engineering maturity, and governance and control the least advanced

AI Engineering Maturity Index: average score per dimension

Average score per index dimension, all organisations (0 to 1) (Q3 2026 n=862)

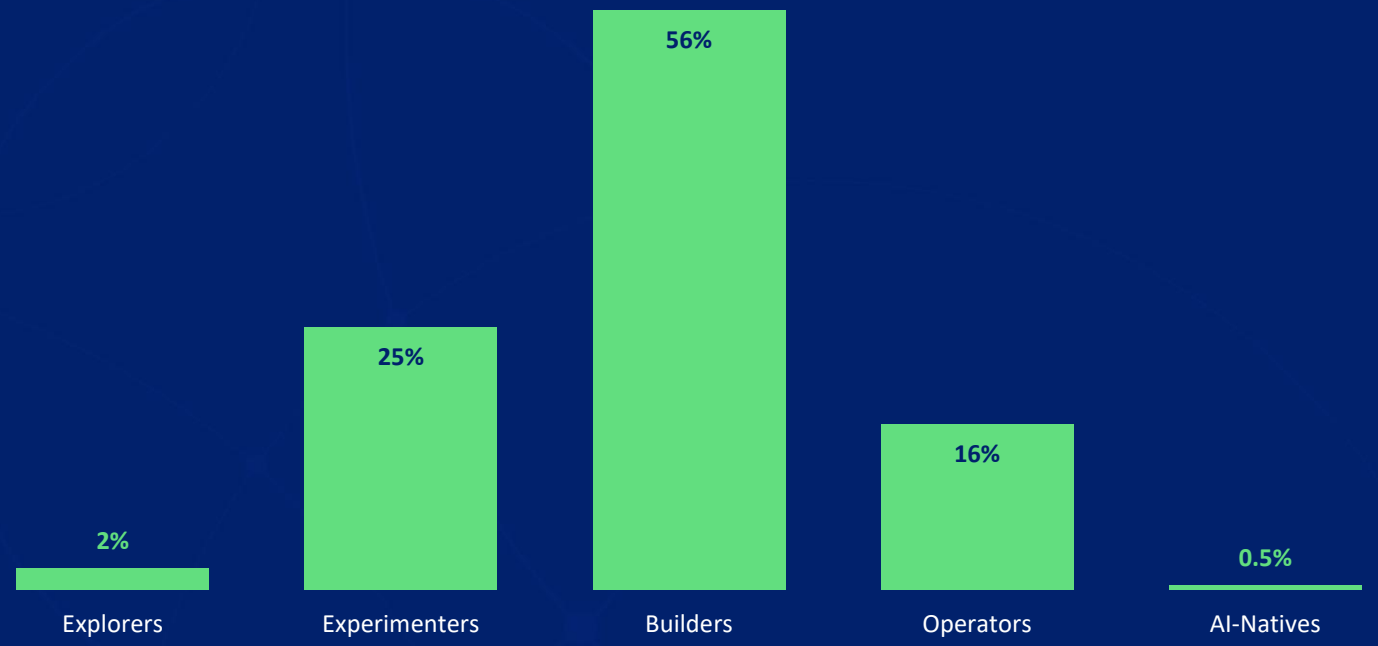


Question wording: Composite score made from 16 questions. See methodology for detailed wording per dimension

More than half of organisations are builders, and experimenters outnumber operators

Distribution of organisations across maturity stages

% of respondents in each maturity stage (Q3 2026 n=862)



Question wording: Composite score made from 16 questions. See methodology for detailed wording per dimension.

Respondents tend to rate their organisation close to where its practices place it, but not evenly across stages. We asked respondents to rate their organisation's AI engineering capability on a five-level scale that mirrors the index's five stages. The most common answer matches the measured stage for builders (44% rate their organisation at level 3, 'Repeatable practices for building and deploying AI') and for operators (46% at level 4, 'Systematic production engineering for AI'). The deviations run in opposite directions. Experimenters and operators tend to rate themselves one level below their measured stage, at 43% and 39% respectively. Builders tend to rate themselves one level above, at 31%. For experimenters, the level below is the most common answer, ahead of the matching level (32%). Organisations' perceptions of their AI engineering maturity do not always align with how systematically their practices are established.

Two readings fit this pattern. Operators may judge themselves against a higher standard because they know more about what systematic AI engineering requires. Builders, who already run AI in production, may equate production with systematic practice. Conversely, the mismatch between observed practices and self-rating could also indicate that they value aspects of maturity not fully captured by the index. Self-ratings are perceptions, and the data cannot separate these explanations.

Experimenters and operators under-rate their organisation, while builders over-rate it

Measured maturity stage against self-rated capability

% of respondents (Q3 2026 n=862)

Computed maturity stage

	Self-rated capability: L1 Individual/ exploratory	L2 Fragmented pilots	L3 Repeatable practices	L4 Systematic production eng.	L5 Org-wide capability
Explorers*	76%	12%	6%	0%	6%
Experimenters	43%	32%	14%	6%	5%
Builders	3%	15%	44%	31%	7%
Operators	2%	1%	39%	46%	12%
AI-Natives*	0%	0%	0%	25%	75%

- >5pp below the average of all other cells within each row
- 2.5 – 5pp below the average of all other cells within each row
- ±2.5pp around the average of all other cells within each row
- 2.5 – 5pp above the average of all other cells within each row
- >5pp above the average of all other cells within each row

***Note:** Explorers and AI-Natives are shown for completeness only, due to small sample size, and treated as directional only.

Question wording: Composite score made from 16 questions. See methodology for detailed wording per dimension.

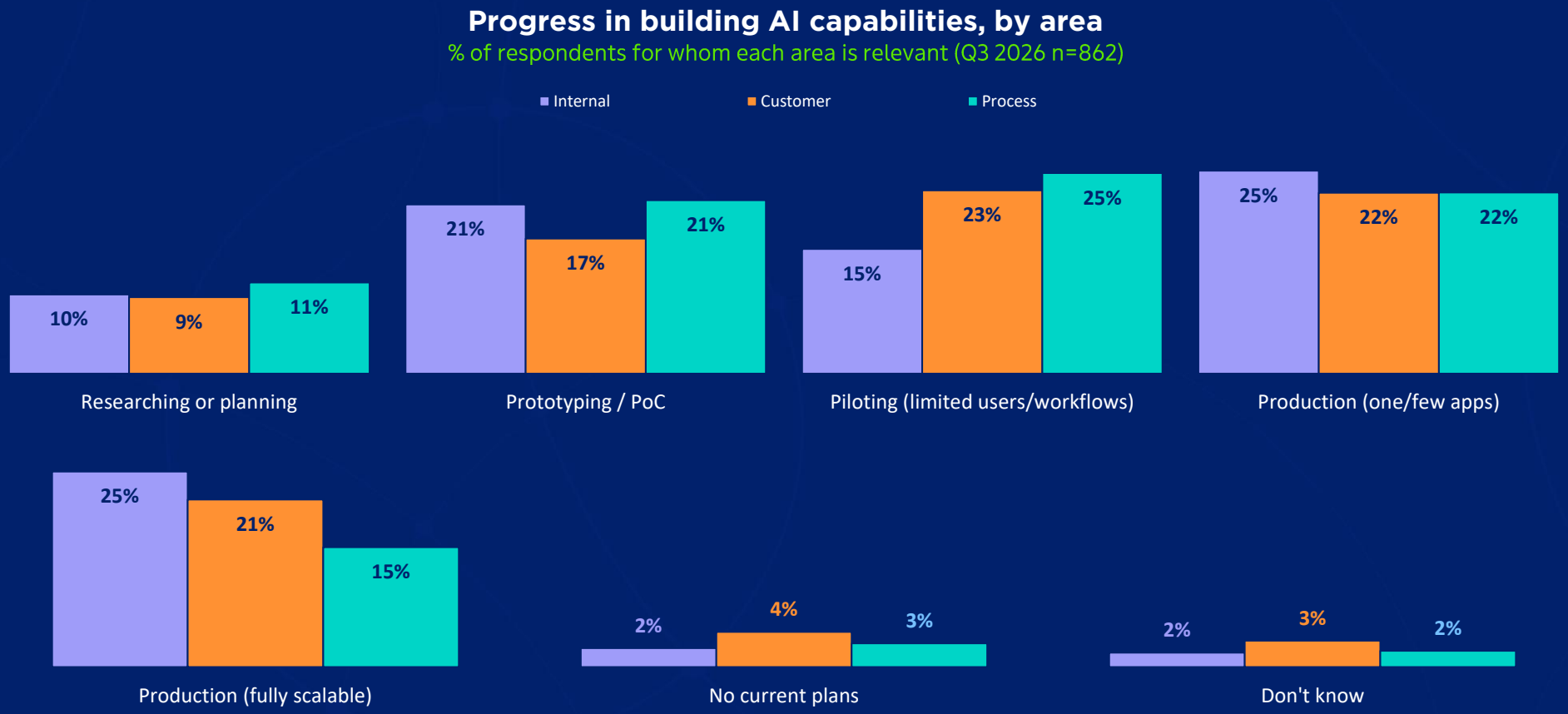
What organisations build, and where

AI capabilities reach production most often in internal software and least often in business processes and workflows. Among organisations for which each area is relevant, half (50%) have AI capabilities live in production in internal software: 25% deployed in one or a few applications and 25% fully scalable to others. The share is 43% for customer software (22% and 21%) and 37% for processes and workflows (22% and 15%). Processes and workflows concentrate instead in piloting (25%) and prototyping (21%).

Business processes and workflows have the highest concentration of organisations in prototyping or piloting and the lowest proportion reporting fully scalable production capabilities.

That could suggest challenges in progressing from experimentation to implementation. A likely explanation is ownership. The engineering team that builds it controls internal software. A business process usually spans several teams, systems, and owners, so scaling AI there requires organisational change as well as technical work.

Half of organisations run AI capabilities in production in internal software



Question wording: For each area, how far has your organisation got in building or integrating AI capabilities?

Organisations mostly build AI capabilities for information tasks. The following figures are among organisations that have reached at least the prototyping stage in each area. Answering questions in natural language is the most common capability in every area: 50% in internal software, 45% in customer software, and 39% in processes and workflows. Beyond that, each area has its own profile.

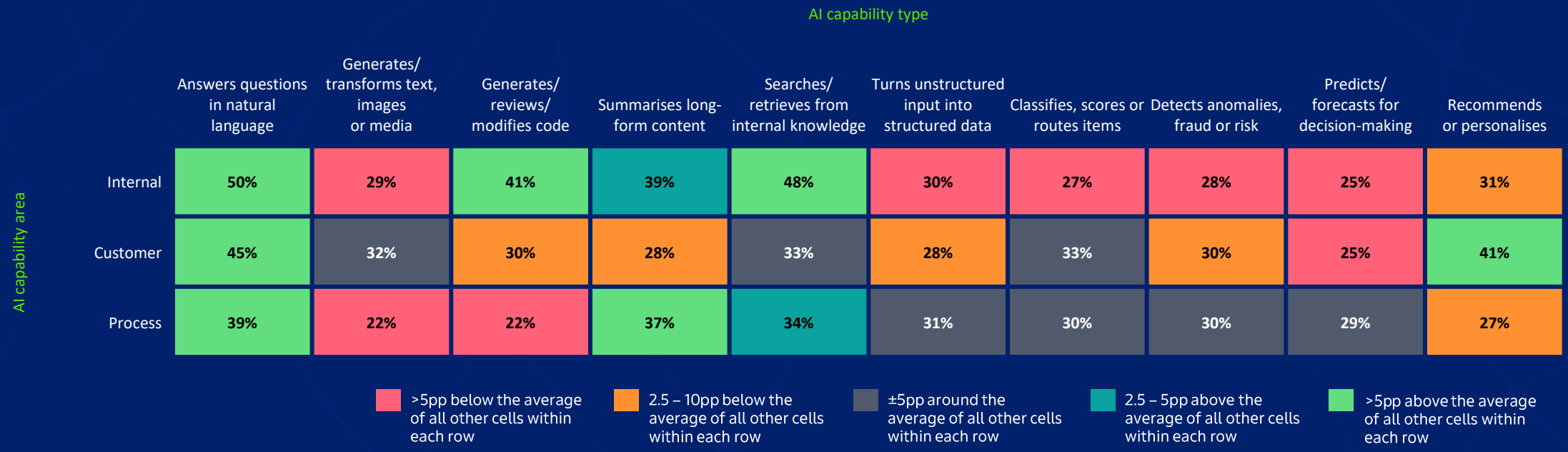
In internal software, searching and retrieving information from internal knowledge bases (48%) and generating, reviewing, or modifying code (41%) follow. In customer software, recommending or personalising options comes second at 41%. In processes and workflows, summarising long-form content (37%) and searching internal knowledge (34%) follow.

Predicting or forecasting values is the least common capability in internal and customer software (25% in each). This may partly reflect the availability of established statistical and machine-learning approaches for numerical prediction, where foundation models are not necessarily the most appropriate choice.

AI capabilities are primarily concentrated around language and information-related activities

Types of AI capabilities built, by deployment area

% of respondents (Q3 2026: internal software n=720, customer software n=527, processes and workflows n=699)



Question wording: For each area, what kinds of AI capabilities has your organisation built or integrated?

Vendor-hosted, off-the-shelf models are the most commonly reported model type in internal software and business workflows, while customer software shows a greater emphasis on fine-tuned models. Among organisations at the prototyping stage or beyond, 63% use vendor-hosted, off-the-shelf models for processes and workflows and 62% for internal software. In customer software, that share falls to 44%, and vendor-hosted fine-tuned models are used just as often (46%, against 30% in internal software and 25% in processes and workflows). Self-hosted open-weight fine-tuned models follow the same pattern: 35% in customer software, against 24% and 18%.

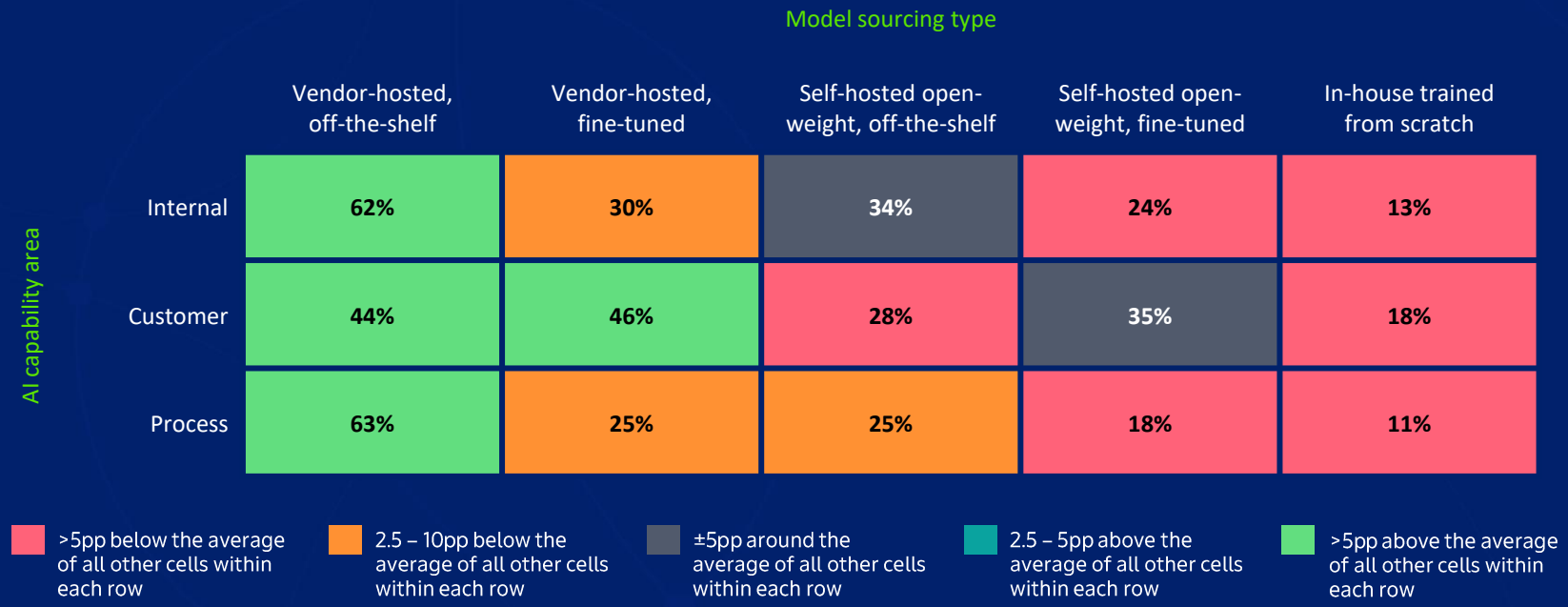
Customer-facing software shows a different model-sourcing profile, with greater use of fine-tuned models across both vendor-hosted and self-hosted approaches. This may reflect stronger requirements for domain-specific performance, specialised functionality or product differentiation. Organisations building AI into commercial software may have greater incentives to adapt models to the particular requirements of their applications, rather than relying exclusively on general-purpose models.

The prevalence of vendor-hosted models highlights the importance of managing external dependencies as AI capabilities become more embedded in organisational software and workflows. Reliance on externally managed models can introduce considerations around service availability, pricing changes, model updates, and continuity of access. These issues become especially consequential when AI capabilities support production applications or business-critical processes. Later in the report, we examine the operational risks organisations have encountered, including the 26% reporting a model change or outage affecting production.

Fine-tuning is more common for customer software, while processes and workflows rely most on vendor-hosted, off-the-shelf models

AI models used by deployment area

% of respondents (Q3 2026: internal software (n=711), customer software (n=518), processes and workflows (n=688)



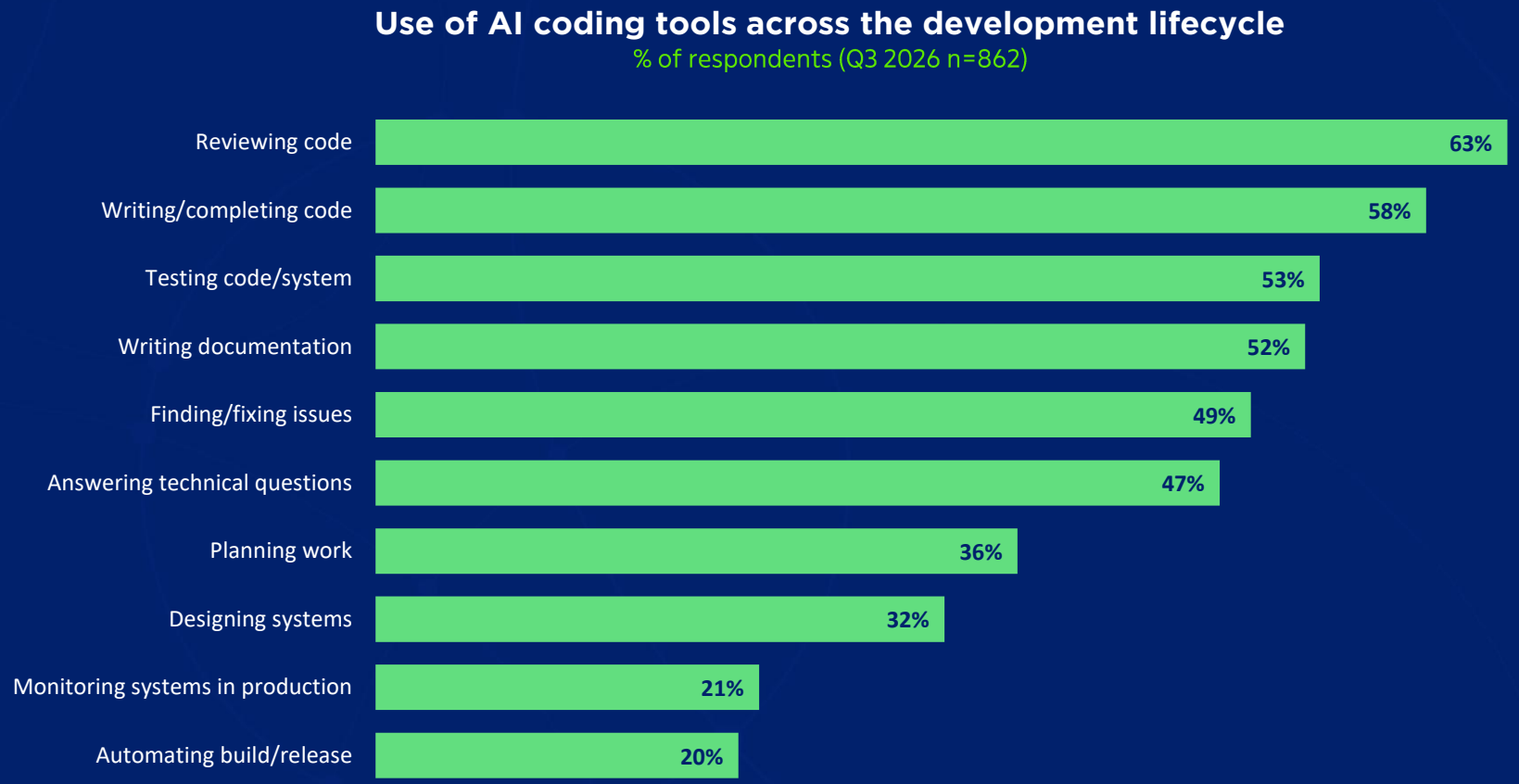
Question wording: For each area, which of the following AI model(s) is used to power the AI capabilities your organisation built or integrated?

AI in the development lifecycle

AI coding tools cluster around code itself and rarely reach operations. 63% of respondents say their team uses AI coding tools to review code, and 58% to write and complete code. Testing (53%) and writing documentation (52%) follow. At the two ends of the lifecycle, use drops. 36% use AI tools for planning work and 32% for designing systems. Only 20% use them to automate build and release, and 21% to monitor systems in production.

Reviewing code ranks above writing it. The prominence of both code review and code generation suggests that AI coding tools are being used not only to produce code but also to support its evaluation and improvement. Their relatively widespread use in testing and debugging reinforces the role these tools play across multiple activities in the coding process. AI coding tools are least used for build, release, and production monitoring.

Teams use AI coding tools most to review and write code, and least to automate build and release or monitor production



Question wording: Which of these activities does your team typically do with help from AI coding tools?

Most organisations have AI in production somewhere, and most of what they build is information work running on vendor models. The practices needed to scale that work, governance above all, lag behind. The next section asks which organisations and practices are associated with higher maturity.

PROFILING MATURE ORGANISATIONS

02

Progress® | /DATA

Which practices are most associated with maturity

To understand what separates leaders from the rest, we examine which specific practices are most strongly associated with higher maturity, spanning architecture, data, safeguards, governance, agents, guidance, and support to engineers.

To compare practices, we calculate each practice's lift: the percentage difference in average AI Engineering Maturity Index scores between developers reporting a practice and those not reporting it. To avoid the direct mechanical relationship between the index and the practice that feeds into it, we exclude the question containing that practice from the index calculation when computing the lift. This lets us assess whether respondents reporting a particular practice also show greater maturity across all other measured AI engineering practices. Hence, a lift of 1% means that respondents reporting the practice have an average maturity score 1% higher than those not reporting it.

Architectural change shows the strongest association with maturity. Practices that change architecture for AI have an average lift of 29%. Systems that let AI capabilities use internal data follow at 27%, and safeguards applied to AI capabilities at 25%. Controls around AI agents (17%), support for engineers developing AI skills (15%), and formal guidance on AI topics (12%) show weaker associations.

The single practice most associated with maturity is an AI gateway or model abstraction layer. 38% of organisations have one, and they score 33% higher on the index than those without one. An AI gateway is a central layer between applications and models. It routes requests to the appropriate model and applies security, access, and cost controls in one place. Data pipelines connecting internal data to AI systems come second (40% adoption, 31% lift). The two are complementary. A gateway controls how applications reach models, and pipelines control how models reach internal data.

A gateway only becomes necessary once an organisation runs several models or applications. Its association with maturity may therefore reflect scale as much as cause it. The data cannot tell which comes first.

Skills and guidance are the practices least associated with maturity. Paid access to AI tools is the most widely adopted practice (64%), but it has a lift of only 2%. Guidance on how AI models work (2%) and on environmental impact (4%) are similarly weak. One hypothesis would be that the most mature organisations hire AI engineers rather than train their software engineers. The data does not support it: hiring people with AI experience has a 16% lift, below average, and 38% of respondents select it. A simpler explanation is diffusion. Skills support has spread to organisations at every stage, so it no longer separates them.

Production safeguards lag furthest behind in their association with maturity. Plotting each practice's lift against its adoption gives four groups, divided at the average of each (20% lift, 36% adoption). ‘Established essentials’ combine above-average lift and adoption. ‘Underused differentiators’ combine above-average lift with below-average adoption. ‘Table stakes’ are widely adopted but weakly associated with maturity, and the remaining ‘niche’ practices are neither.

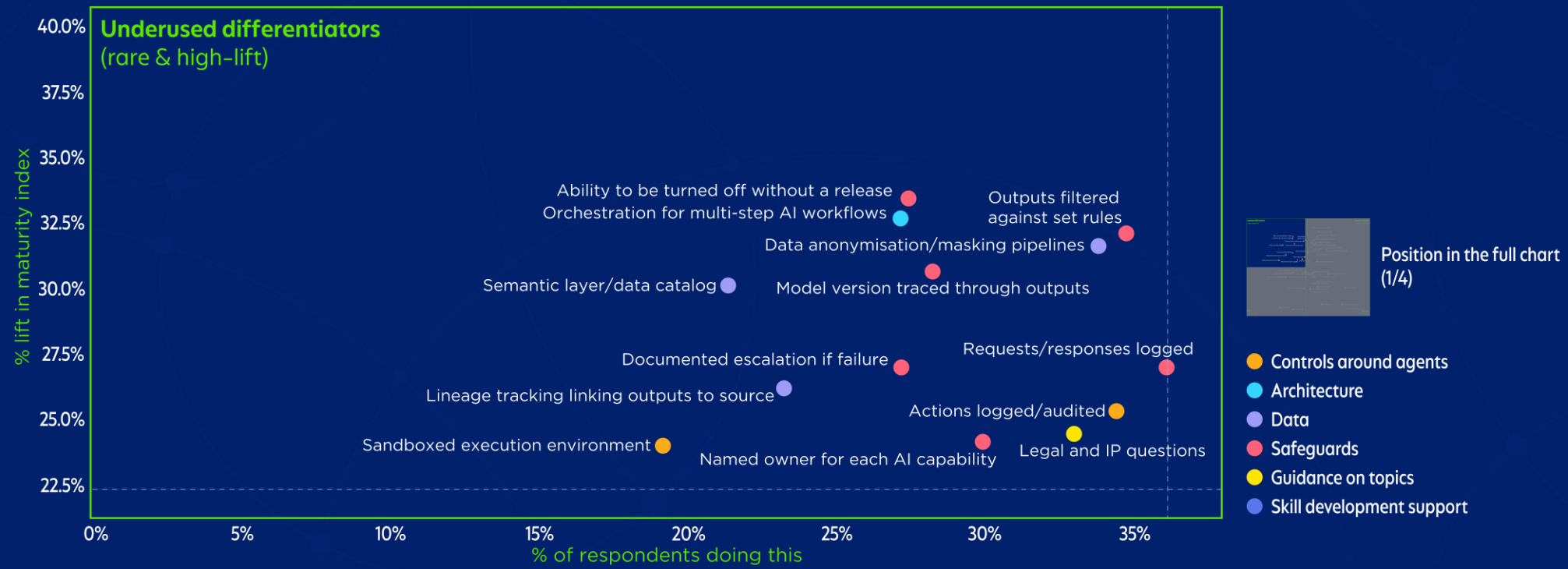
Six of the eight safeguards fall among the underused differentiators, against one of six architectural practices and three of seven data practices. They are the ability to turn an AI capability off without a release, outputs filtered against set rules, model version traced through outputs, requests and responses logged, documented escalation if failure, and a named owner for each AI capability. Each is adopted by 27% to 36% of organisations, and each has a lift of 21% to 31%.

These safeguards decide how an organisation responds when an AI capability in production fails. They belong to the governance and control dimension, the least advanced of the five. *Making them standard practice is likely to be the most direct route to higher maturity for most organisations.*

Six of the eight production safeguards are underused differentiators: strongly associated with maturity, but adopted by a minority of organisations

Adoption against association with maturity, by practice

% of respondents adopting each practice against its lift in maturity index. Dashed lines mark the average on each axis (Q3 2026 n=862)

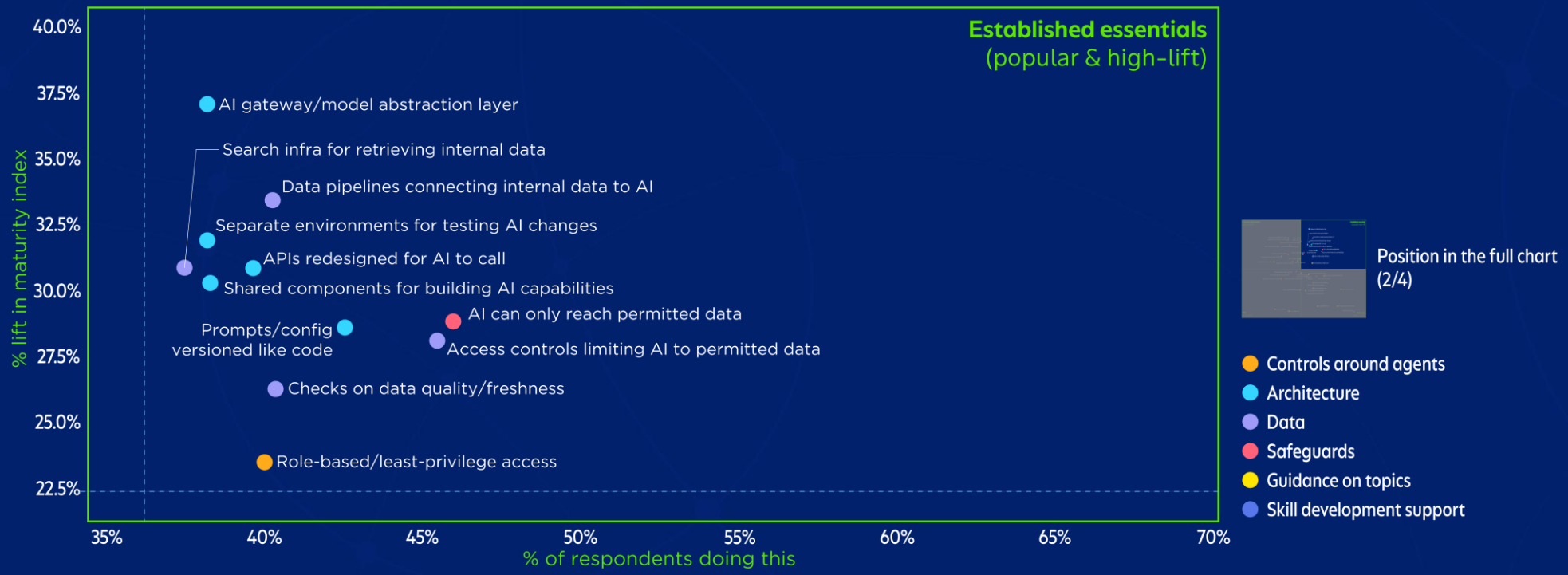


Question wording: See methodology.

Six of the eight production safeguards are underused differentiators: strongly associated with maturity, but adopted by a minority of organisations

Adoption against association with maturity, by practice

% of respondents adopting each practice against its lift in maturity index. Dashed lines mark the average on each axis (Q3 2026 n=862)

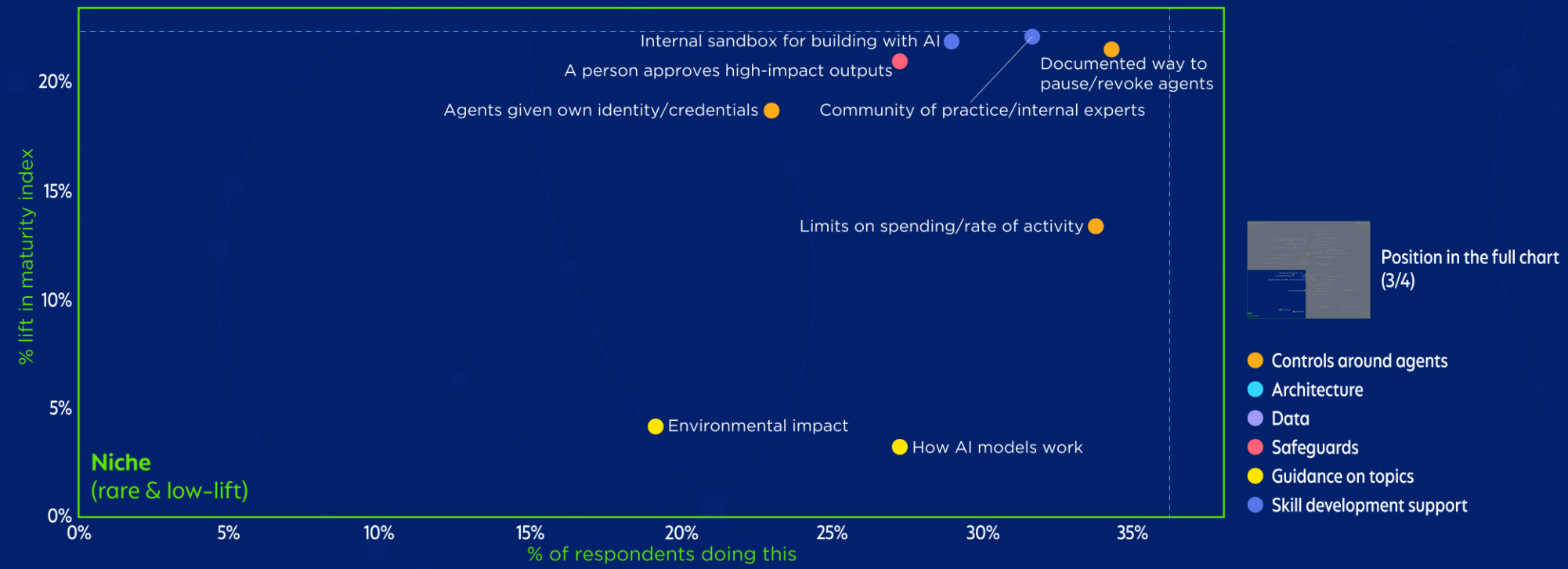


Question wording: See methodology.

Six of the eight production safeguards are underused differentiators: strongly associated with maturity, but adopted by a minority of organisations

Adoption against association with maturity, by practice

% of respondents adopting each practice against its lift in maturity index. Dashed lines mark the average on each axis (Q3 2026 n=862)

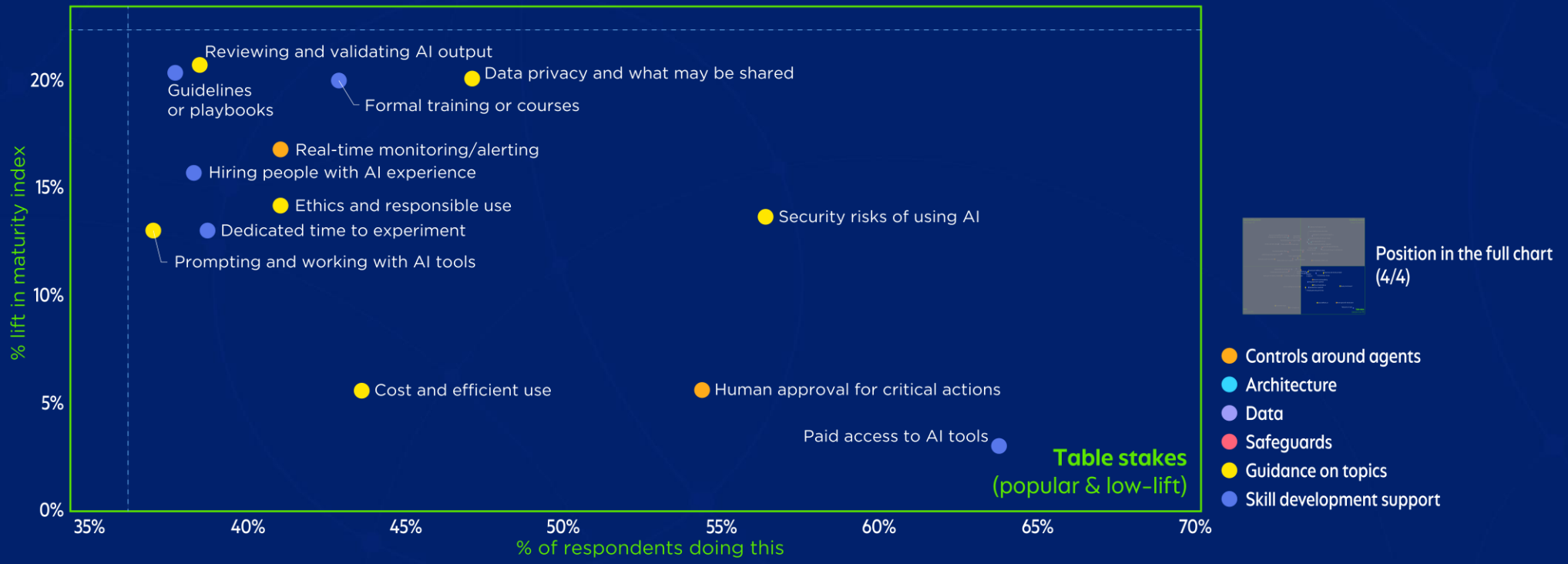


Question wording: See methodology.

Six of the eight production safeguards are underused differentiators: strongly associated with maturity, but adopted by a minority of organisations

Adoption against association with maturity, by practice

% of respondents adopting each practice against its lift in maturity index. Dashed lines mark the average on each axis (Q3 2026 n=862)



Question wording: See methodology.

Which organisations are most mature

The larger the organisation, the more mature it tends to be. 28% of organisations with 1,001 to 10,000 employees are operators, compared with 16% on average, and 17% are experimenters, compared with 25% on average. Among organisations with 21 to 100 employees, 4% are operators, 11% are explorers, and 42% are experimenters. Organisations with 101 to 500 employees also lean towards experimenters (30%).

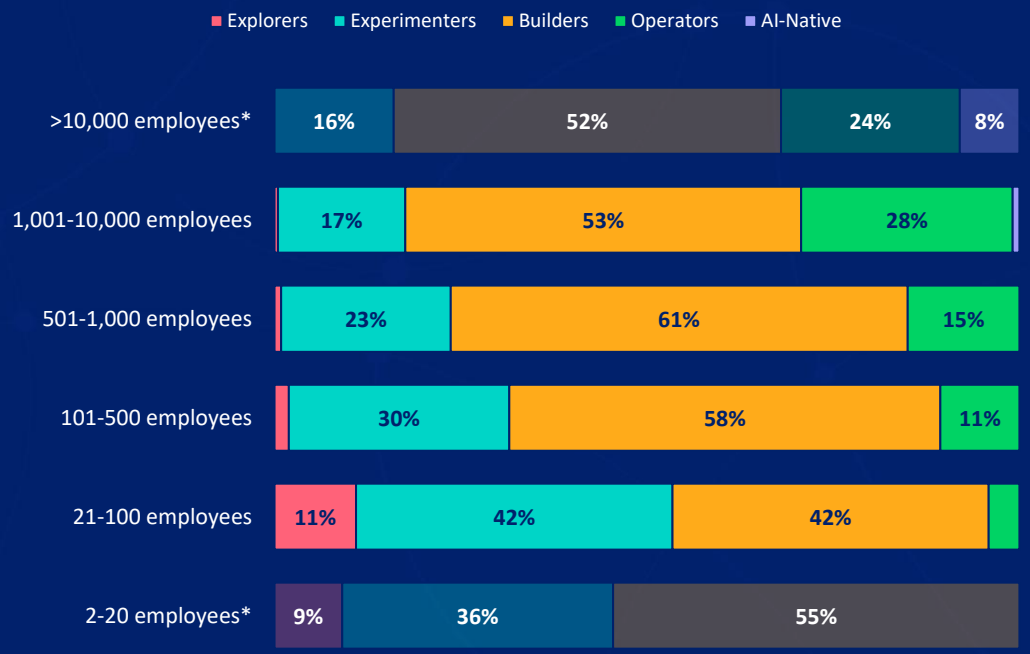
AI engineering maturity shows a clear positive association with organisation size. Smaller organisations are more concentrated in the experimentation stages, while larger organisations are considerably more likely to reach systematic production engineering. Larger organisations are more likely to have the infrastructure, budget, and specialist headcount that scaling AI requires. A second explanation is also possible: respondents from larger organisations may hold more specialised AI roles. The two cannot be separated with this data.

Organisation age shows no comparable relationship with maturity. Organisations more than 30 years old are the most likely to be experimenters (35%), yet they also match the highest share of operators (18%, level with organisations aged 16 to 30 years).

AI engineering maturity increases with organisation size, but shows a less consistent relationship with organisation age

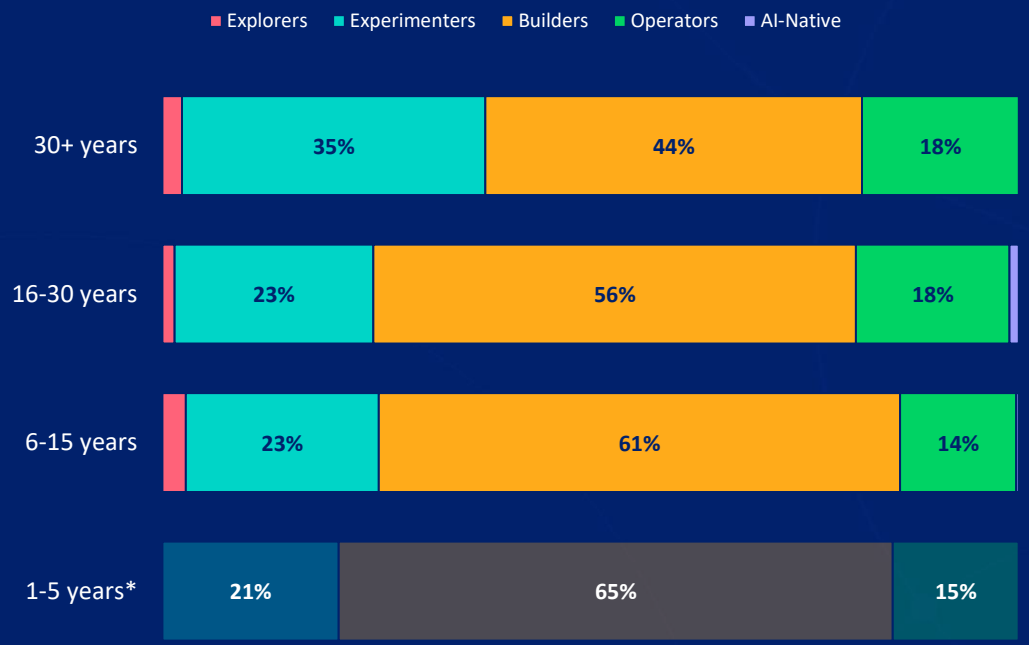
Maturity stage by organisation size

% of respondents in each maturity stage (Q3 2026 n=862)



Maturity stage by organisation age

% of respondents in each maturity stage (Q3 2026 n=862)



***Note:** >10,000 employees, 2-20 employees and 1-5 years are shown for completeness only, due to small sample size, and treated as directional only.

Question wording: How many employees does the organisation that you work for have worldwide? | How many years has your organisation been in operation? | Composite score. See methodology

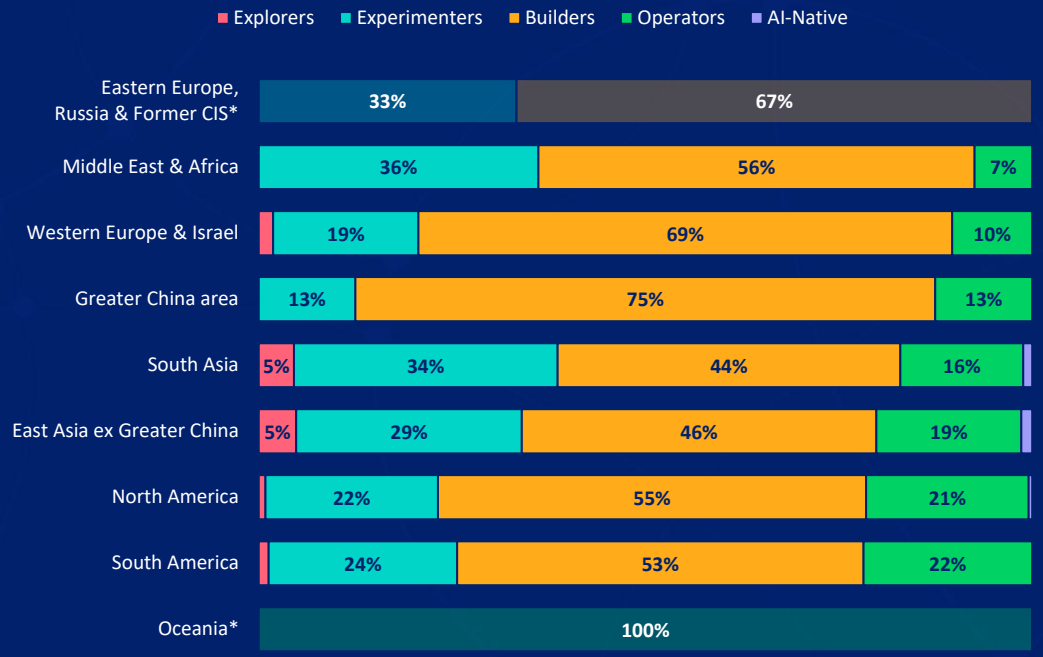
Maturity also varies by region. South America (22%), North America (21%), and East Asia excluding Greater China (19%) have the highest shares of operators, compared with an average of 16%. Western Europe stands apart, as 69% of its organisations are builders, the highest share of any region, but only 10% are operators. South Asia (34%) and the Middle East and Africa (36%) have the highest shares of experimenters, against 25% on average.

Maturity varies by industry too. Software, SaaS, and technology products (20%), cloud and telecom infrastructure (20%), and financial services (18%) have the highest shares of operators. The first two are the most technology-centred industries, which would let them scale AI earlier. Financial services has long invested in data and automation, and fintech growth is likely to reinforce that. At the other end, experimenters are over-represented in consumer, retail, and media (41%), healthcare and life sciences (38%), industry, energy, and mobility (32%), and professional and business services (28%).

Organisations in North and South America and in technology and financial services are the most likely to be operators

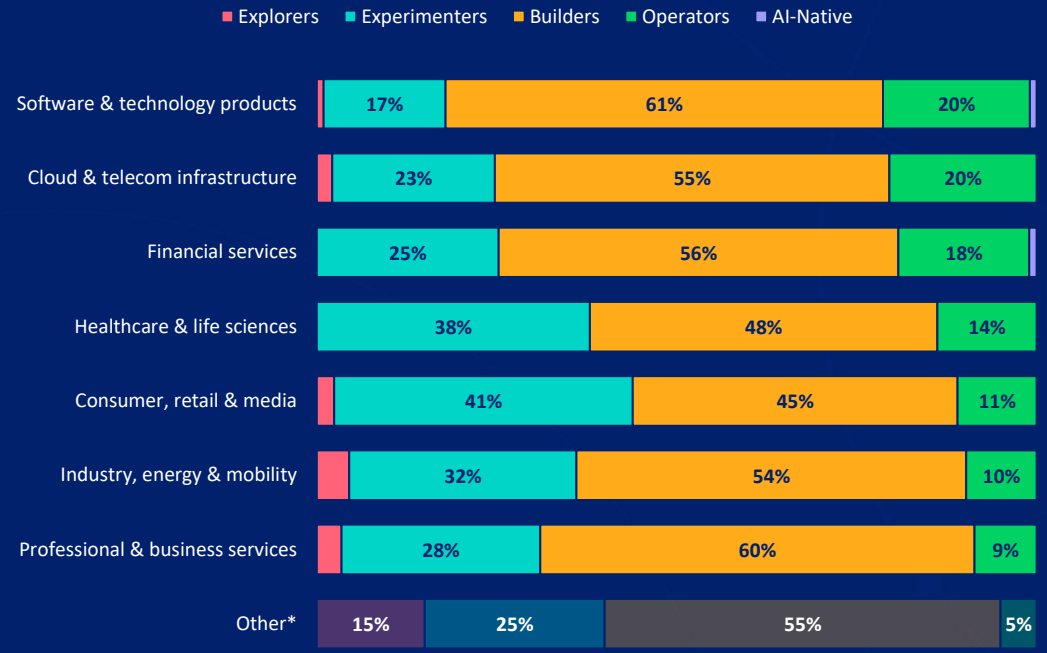
Maturity stage by region

% of respondents in each maturity stage. Regions with too few respondents shaded (Q3 2026 n=862)



Maturity stage by industry

% of respondents in each maturity stage. Industries with too few respondents shaded (Q3 2026 n=862)



***Note:** Eastern Europe, Russia & Former CIS, Oceania and Other are shown for completeness only, due to small sample size, and treated as directional only.

Question wording: Which country are you based in? | Which of the following best describes the main industry your organisation operates in? | Composite score. See methodology

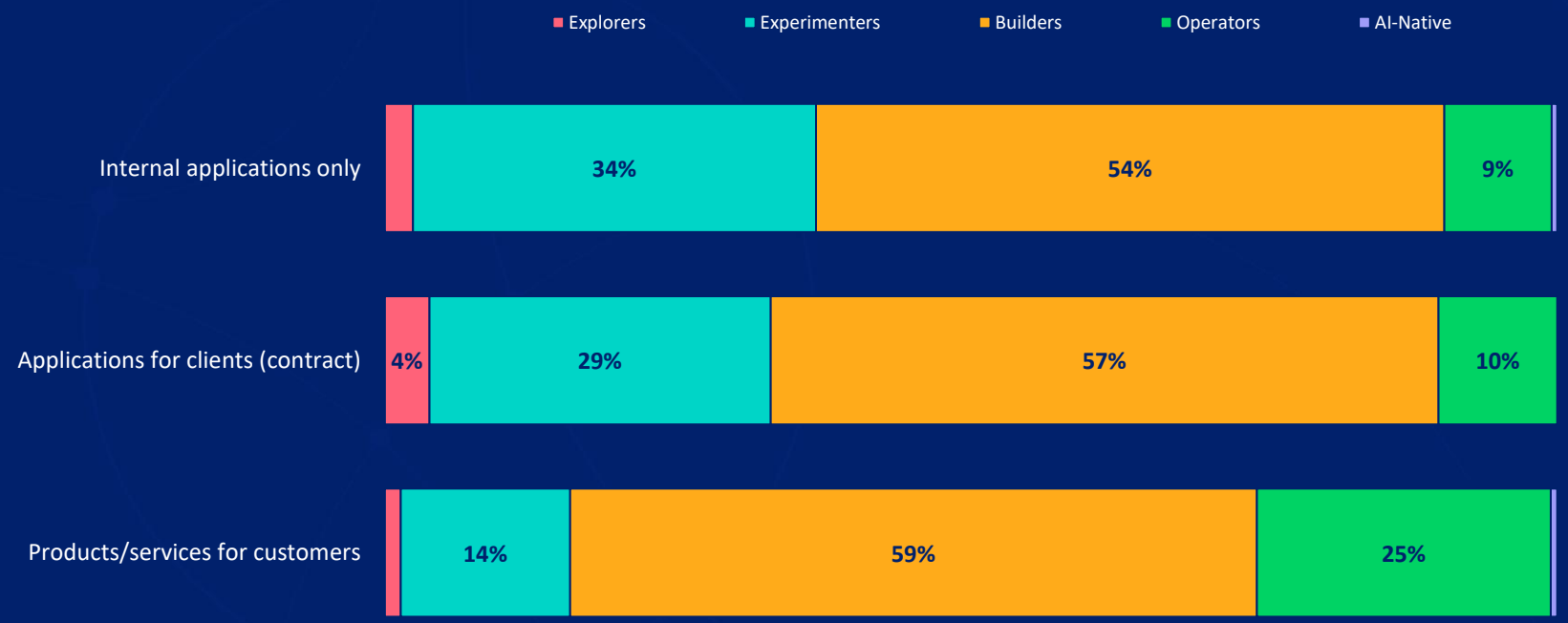
An organisation's relationship to software also appears closely associated with AI engineering maturity. Among respondents working for organisations that develop or operate software products and services for customers, 25% are classified as operators, compared with 10% among those developing software on behalf of clients and 9% among those building applications primarily for internal use.

One possible explanation is the greater emphasis on sustained product delivery and operational responsibility among software product companies. When AI capabilities form part of commercial products, organisations may face stronger requirements for reliability, ongoing monitoring, release management, and consistent engineering practices. However, these differences may also reflect variation in organisation size, technical specialisation, and the extent to which AI is central to the business.

Organisations building software products for customers are more than twice as likely to reach the operator stage as those developing internal or client applications

Maturity stage by organisation's relationship to software

% of respondents in each maturity stage (Q3 2026 n=862)



Question wording: Which best describes your organisation's relationship to software? | Composite score. See methodology

KEEPING PRODUCTION AI UNDER CONTROL

03

Progress® | /DATA

As agent autonomy increases, controls rise – but so do unintended actions

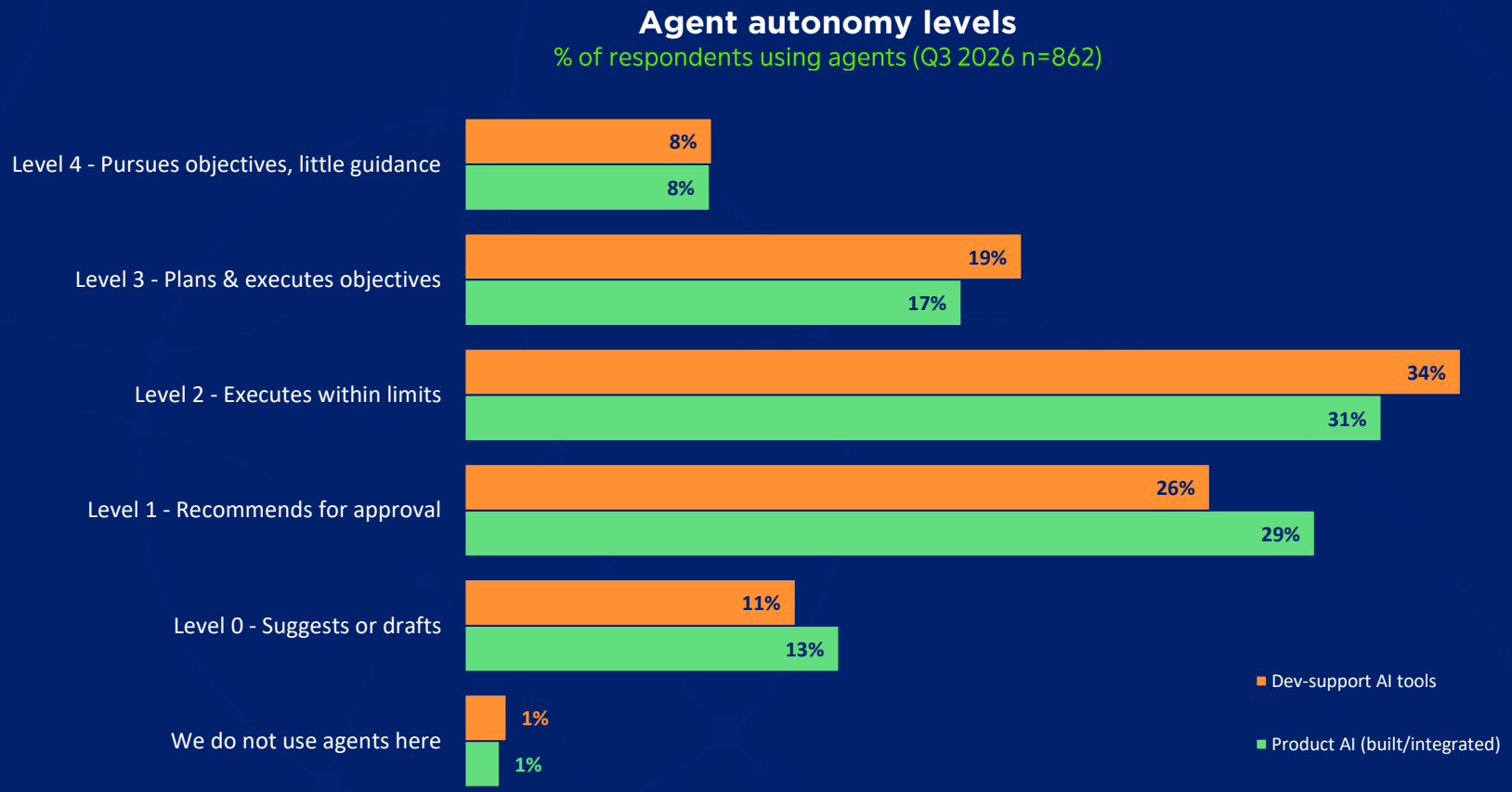
Substantial agent autonomy is already permitted in both product AI and in developer support AI tools. Furthermore, organisations give AI agents slightly more autonomy in their own development work than in the products they build. Among organisations that use agents, 19% allow agents in their development tools to plan and execute assigned objectives, compared with 17% for agents in the AI features they build. 34% allow development agents to execute assigned tasks within limits, against 31% for product agents.

The number of reported governance controls increases notably once agents are permitted to execute actions. Organisations whose agents only suggest or draft have an average of 1.7 controls in place and 2.1 where agents recommend actions for approval. The average rises to 3.3 when agents execute tasks within limits and 3.2 when they plan and execute objectives. Where agents pursue objectives with little guidance, it falls back to 2.7.

Organisations that let agents execute actions report unintended agent actions more often than those that limit agents to suggestions or recommendations, although the pattern is not consistent at the highest autonomy level. 7% of organisations whose agents suggest or recommend report that an agent took an unintended action. That share rises to 18% where agents execute within limits and 25% where they plan and execute objectives. At the highest level of autonomy, it falls to 12%.

The highest level combines fewer controls with fewer reported incidents. Two explanations are possible. These organisations may be more experienced with agents and deploy them more carefully. Alternatively, fewer controls such as logging and monitoring may mean fewer incidents are detected in the first place. The data cannot separate the two, and the finding deserves closer study.

Around one-third permit AI agents to execute tasks within defined limits in both development-support tools and product AI

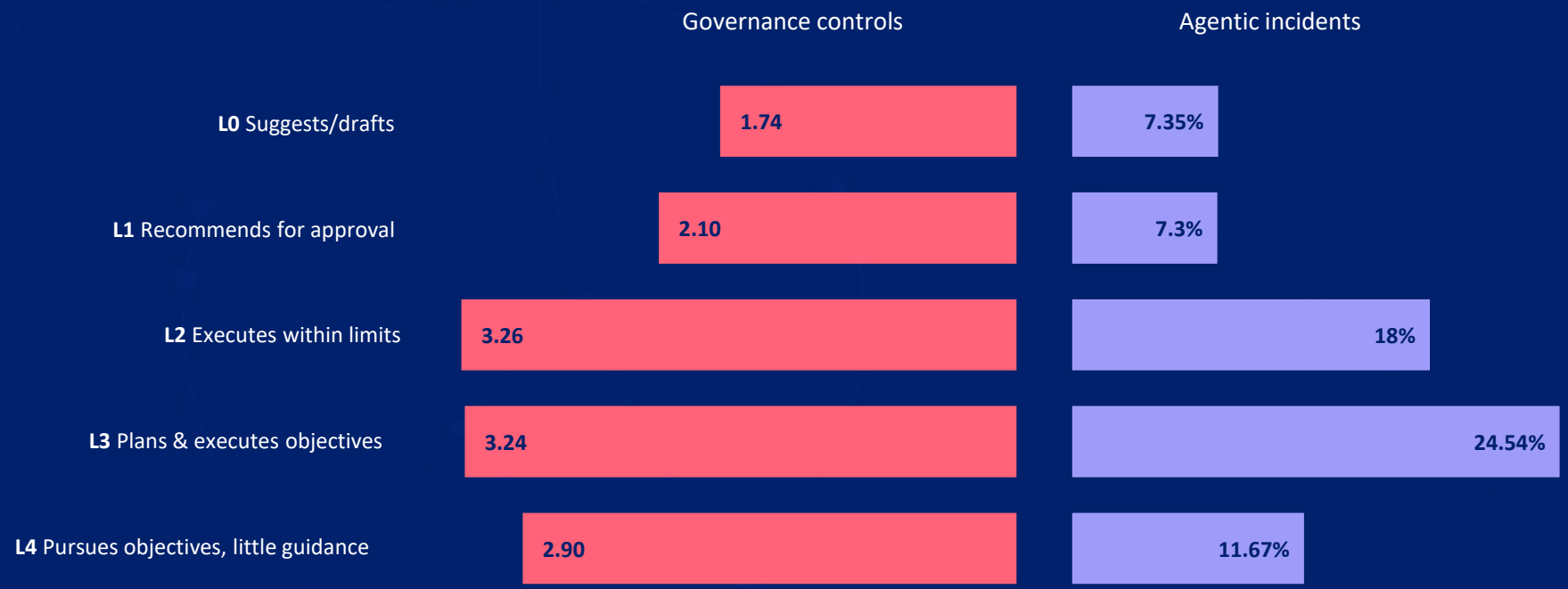


Question wording: What is the highest level of autonomy AI agents are permitted in each area?

Organisations with the highest level of agent autonomy reported fewer controls and fewer incidents

Agent autonomy against controls and unintended actions

Average number of controls and % reporting an unintended agent action, by highest autonomy level (Q3 2026 n=841)



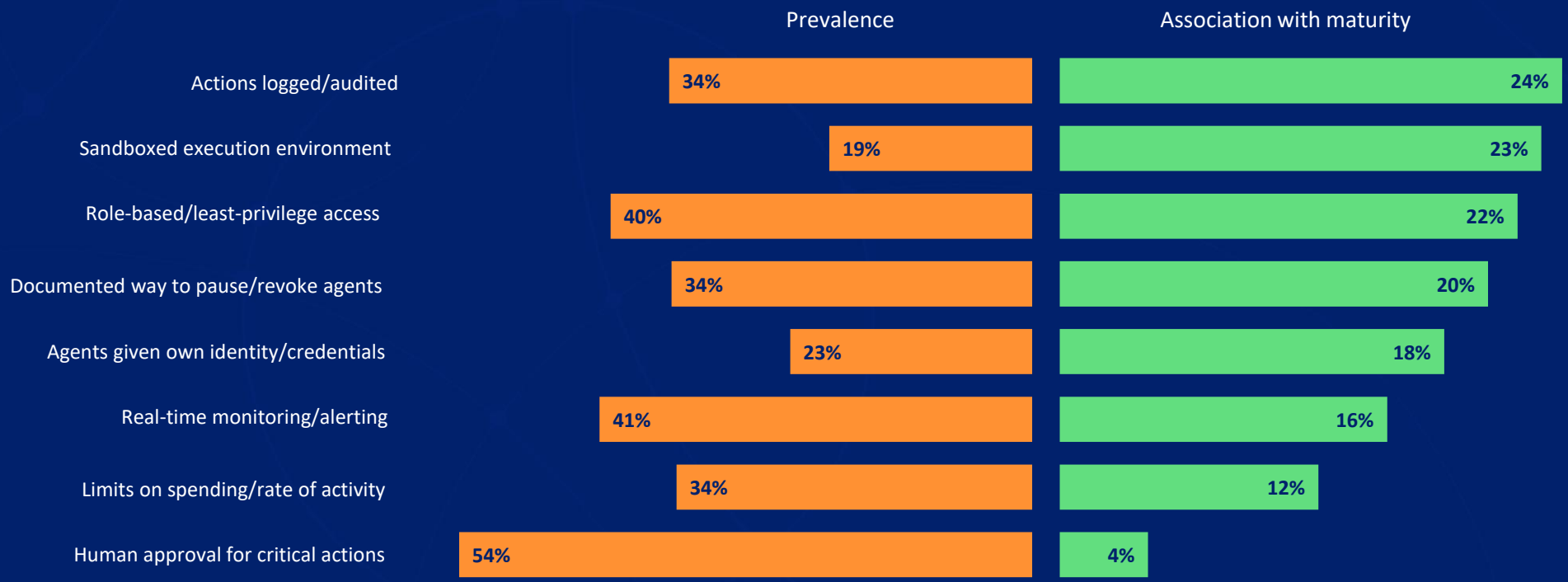
Question wording: What is the highest level of autonomy AI agents are permitted in each area? | Which of the following controls, if any, apply to AI agents in your organisation? | Which of the following has your organisation experienced in the past 12 months?

Human approval is the most common agent control but the least associated with maturity. 54% of organisations require human approval for critical agent actions, more than any other control. Its lift is 4%, the weakest of all controls. Logging and auditing agent actions (lift of 24%), sandboxed execution environment (lift of 23%), and role-based or least-privilege access (lift of 22%) are the controls most associated with maturity. Human approval is a default that organisations adopt at every stage. Mature organisations add structural controls that don't depend on a person reviewing each action, which will likely matter more as the number of agent actions grows.

Human approval is the most commonly reported agent control, while logging, sandboxing, and least-privilege access show higher maturity lift

Agent autonomy against controls and unintended actions

% of respondents | lift (Q3 2026 n=862)



Question wording: Which of the following controls, if any, apply to AI agents in your organisation?

Evaluating before release, monitoring after it

For most organisations, rigour before release and rigour after it move together. We compared each organisation's level on a before-release evaluation ladder with its level on an after-release monitoring ladder, for organisations with AI at piloting stage or beyond. Both ladders run from no defined process or very informal checks to systematically automated. Levels match for 40% of organisations in customer software, 36% in internal software, and 32% in processes and workflows.

Where they differ, the direction depends on the area. In internal software, evaluation more often leads monitoring (36%) than follows it (28%). In customer software, the reverse holds: monitoring is ahead for 32% and evaluation for 28%. Processes and workflows follow the same pattern, with monitoring ahead at 35% and evaluation at 33%.

Processes and workflows show a gap at the top of the ladder. Among organisations whose evaluation can automatically block a release, 49% in internal software and 58% in customer software also have monitoring that can trigger a fallback or rollback. In processes and workflows, that share is 11%, and 18% only review AI behaviour when problems are reported. The base for this comparison is small, so the finding is directional. Still, it raises the question of whether automated recovery mechanisms are keeping pace with more advanced evaluation practices in workflow environments.

Pre-release evaluation and post-release monitoring are not consistently developed to the same relative level across AI deployment areas

Before-release evaluation against after-release monitoring, by area

% of respondents at each monitoring level (Q3 2026 internal software n=542, customer software n=412, processes and workflows n=515)

Internal software

Before release \ After release	*No monitoring	Reviews when reported	Logs on demand	Defined measures tracked	Automated alerts	Can trigger fallback/ rollback
*No defined process	13%	20%	20%	20%	13%	13%
Informal/case-by-case	10%	24%	41%	10%	7%	7%
Manual (documented)	2%	13%	35%	31%	15%	4%
Manual (pass/fail)	1%	6%	28%	39%	24%	2%
Automated at release	1%	5%	10%	45%	34%	6%
Automated (can block)	3%	0%	6%	10%	32%	49%

Customer software

Before release \ After release	*No monitoring	Reviews when reported	Logs on demand	Defined measures tracked	Automated alerts	Can trigger fallback/ rollback
*No defined process	17%	33%	0%	17%	17%	17%
Informal/case-by-case	2%	34%	29%	15%	12%	7%
Manual (documented)	4%	11%	30%	27%	20%	8%
Manual (pass/fail)	1%	7%	19%	26%	32%	15%
Automated at release	3%	10%	14%	14%	49%	9%
Automated (can block)	0%	2%	7%	4%	29%	58%

Processes & workflows

Before release \ After release	*No monitoring	Reviews when reported	Logs on demand	Defined measures tracked	Automated alerts	*Can trigger fallback/ rollback
*No defined process	11%	44%	0%	11%	11%	22%
Informal/case-by-case	5%	29%	45%	8%	8%	6%
Manual (documented)	2%	14%	37%	26%	19%	3%
Manual (pass/fail)	1%	8%	24%	36%	27%	5%
Automated at release	0%	7%	14%	36%	33%	10%
Automated (can block)	5%	18%	11%	27%	27%	11%

- >5pp below the average of all other cells within each row
- 2.5 – 10pp below the average of all other cells within each row
- ±5pp around the average of all other cells within each row
- 2.5 – 5pp above the average of all other cells within each row
- >5pp above the average of all other cells within each row

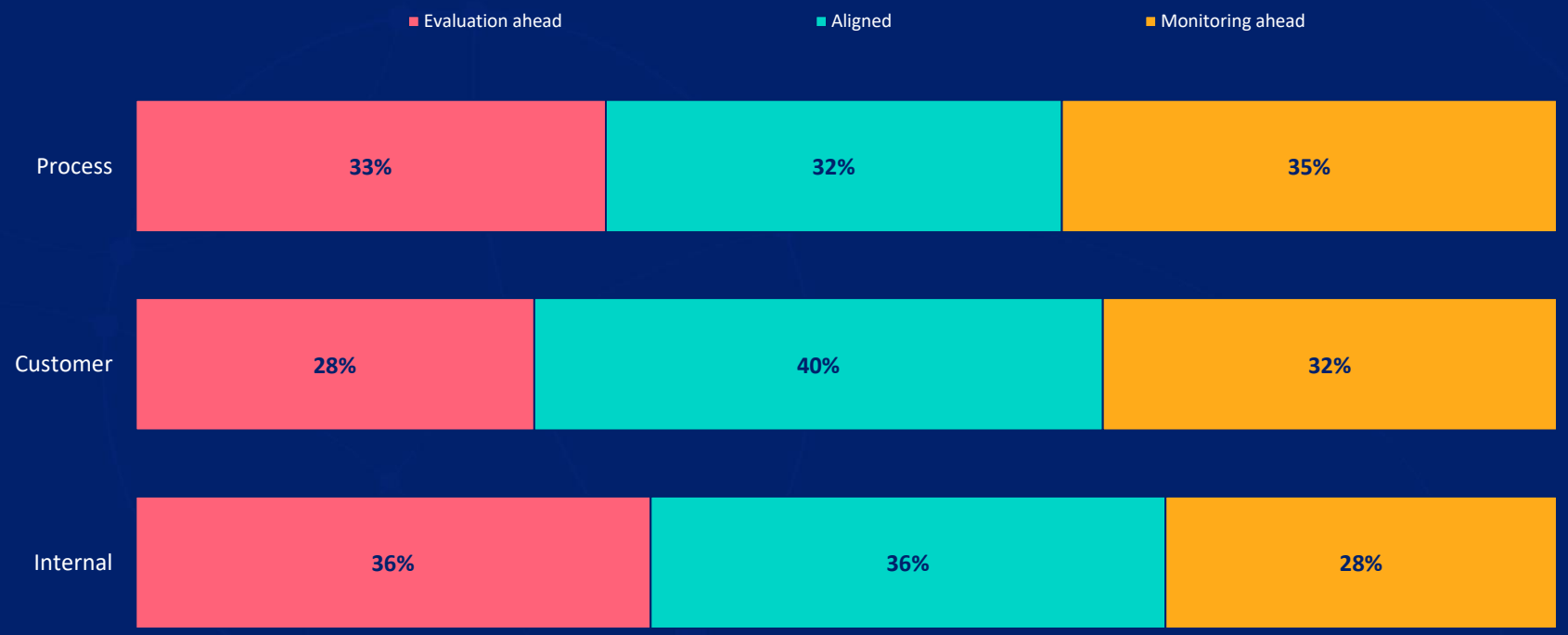
***Note:** No monitoring, no defined process, and can trigger fallback/rollback for processes and workflows, are shown for completion only, due to small sample size, and treated as directional only.

Question wording: For each area, before release, which best describes the evaluation process typically applied to AI capabilities your organisation builds or integrates? | For each area, after release, which best describes the monitoring process typically applied to AI capabilities your organisation builds or integrates?

Evaluation and monitoring keep pace with each other in internal and customer software but less so in processes and workflows

Before-release evaluation against after-release monitoring, by area

% of respondents (Q3 2026 n=862)



Question wording: For each area, before release, which best describes the evaluation process typically applied to AI capabilities your organisation builds or integrates? | For each area, after release, which best describes the monitoring process typically applied to AI capabilities your organisation builds or integrates?

Where AI incidents happen

AI incidents arise both when organisations develop software with AI and when they run AI capabilities in production. 34% of respondents say AI costs significantly exceeded expectations in the past 12 months, and 33% say AI-generated code caused a security or quality issue. A model change or outage hit production for 26%, and incorrect AI output reached users for 25%. The two most common incidents concern what AI costs and the code it helps write. The next two concern AI capabilities once they are live.

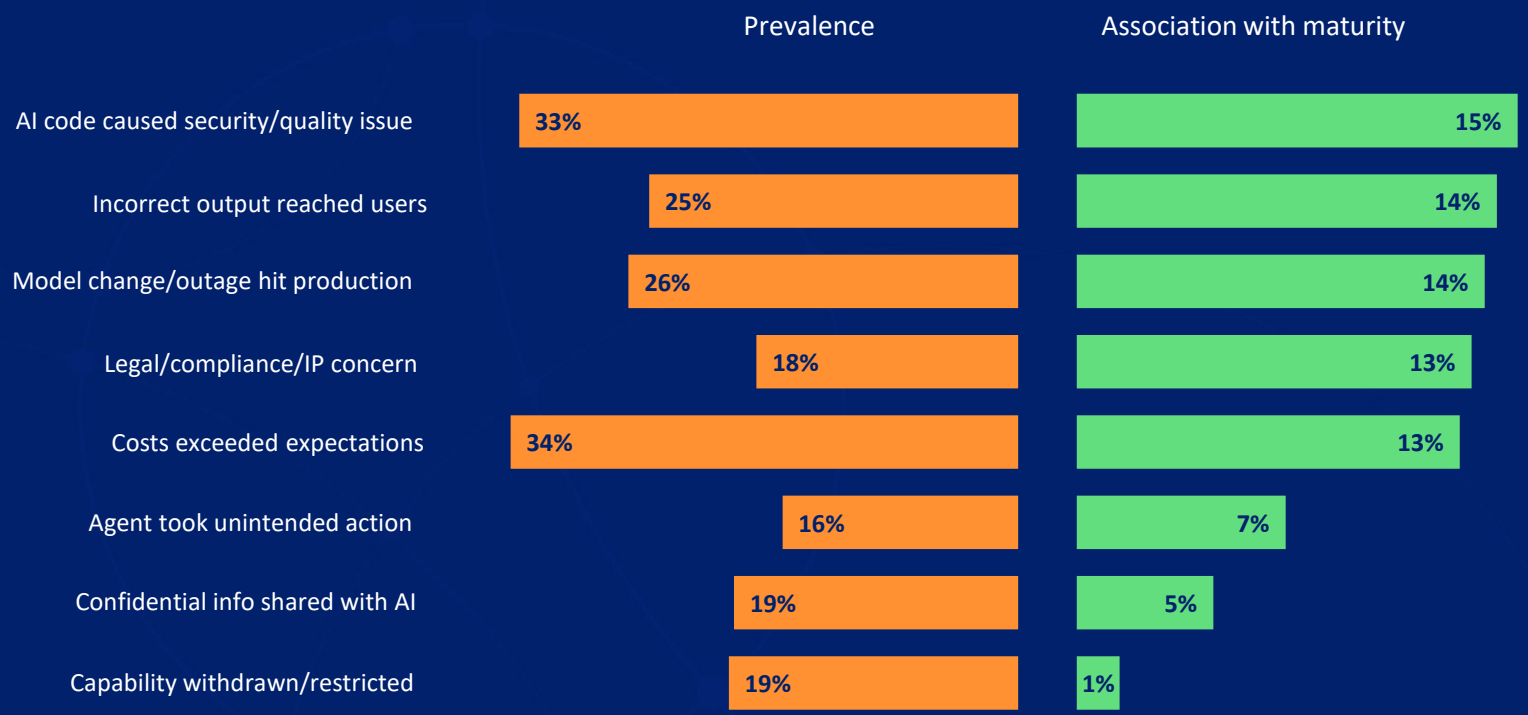
Individual incidents are only weakly associated with maturity. Here, a positive lift means that organisations that experienced the incident score higher on the maturity index than those that did not. Across the eight incidents, the average lift is 10%, against 22% for the practices examined in chapter two. Three incidents sit well below that average: an AI capability withdrawn or restricted (1%), confidential information shared with an AI service (5%), and an AI agent taking an unintended action (7%).

The number of incidents, by contrast, rises clearly with maturity. Operators report an average of 2.4 incident types, builders 2.1, and experimenters 1.2. Two explanations are likely to combine. Mature organisations run more AI in production, so they are more exposed. They also monitor more, so they detect more of what goes wrong. The data cannot separate exposure from detection.

Unexpected AI costs and AI-generated code issues are the most commonly reported AI-related incidents

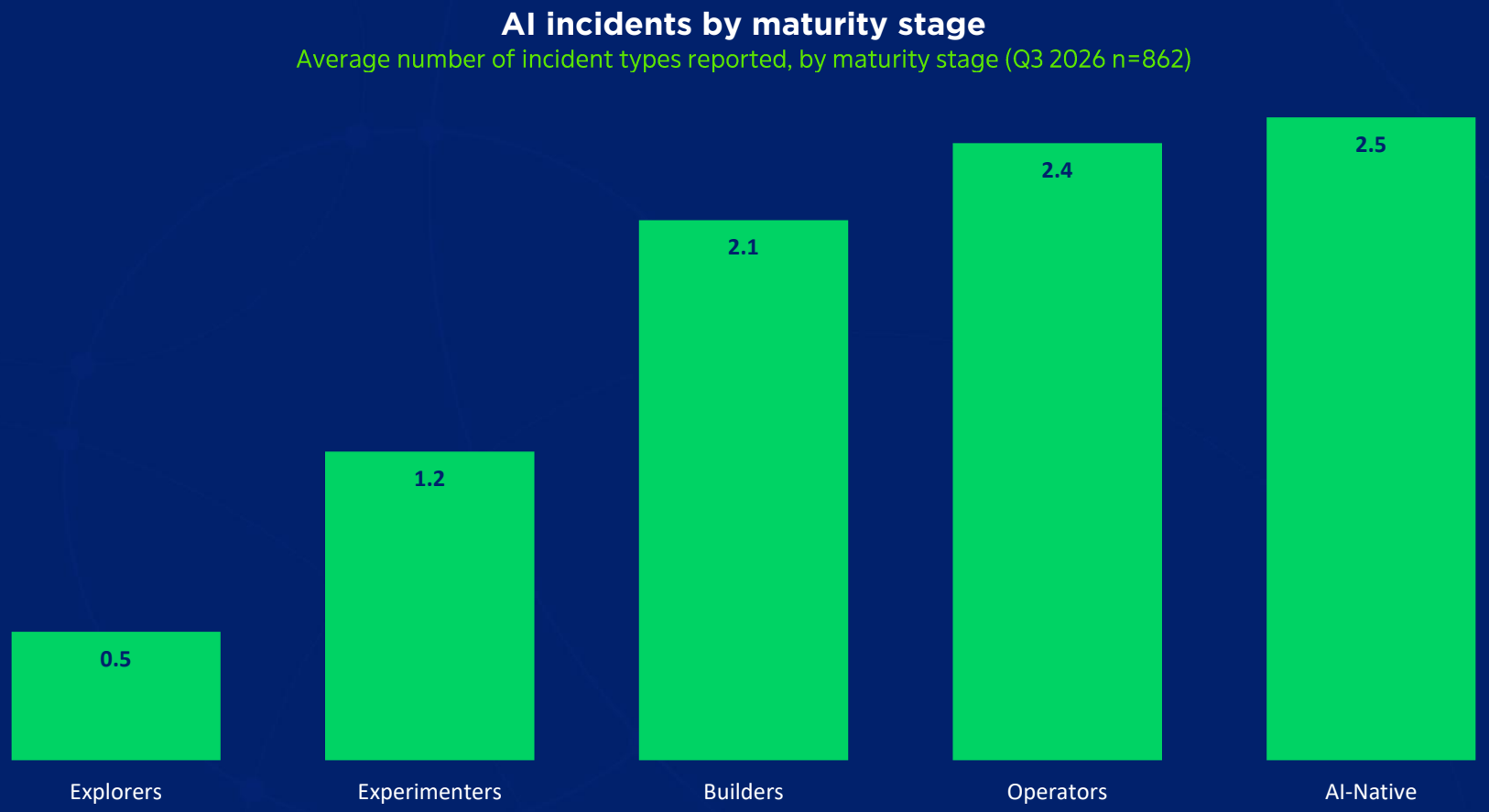
AI incidents in the past 12 months

% of respondents experiencing each incident, and its lift in maturity index (Q3 2026 n=862)



Question wording: Which of the following has your organisation experienced in the past 12 months?

More mature organisations report more AI-related incidents



Question wording: Which of the following has your organisation experienced in the past 12 months?

THE FUTURE OF SOFTWARE DEVELOPMENT

04

Progress® | /DATA

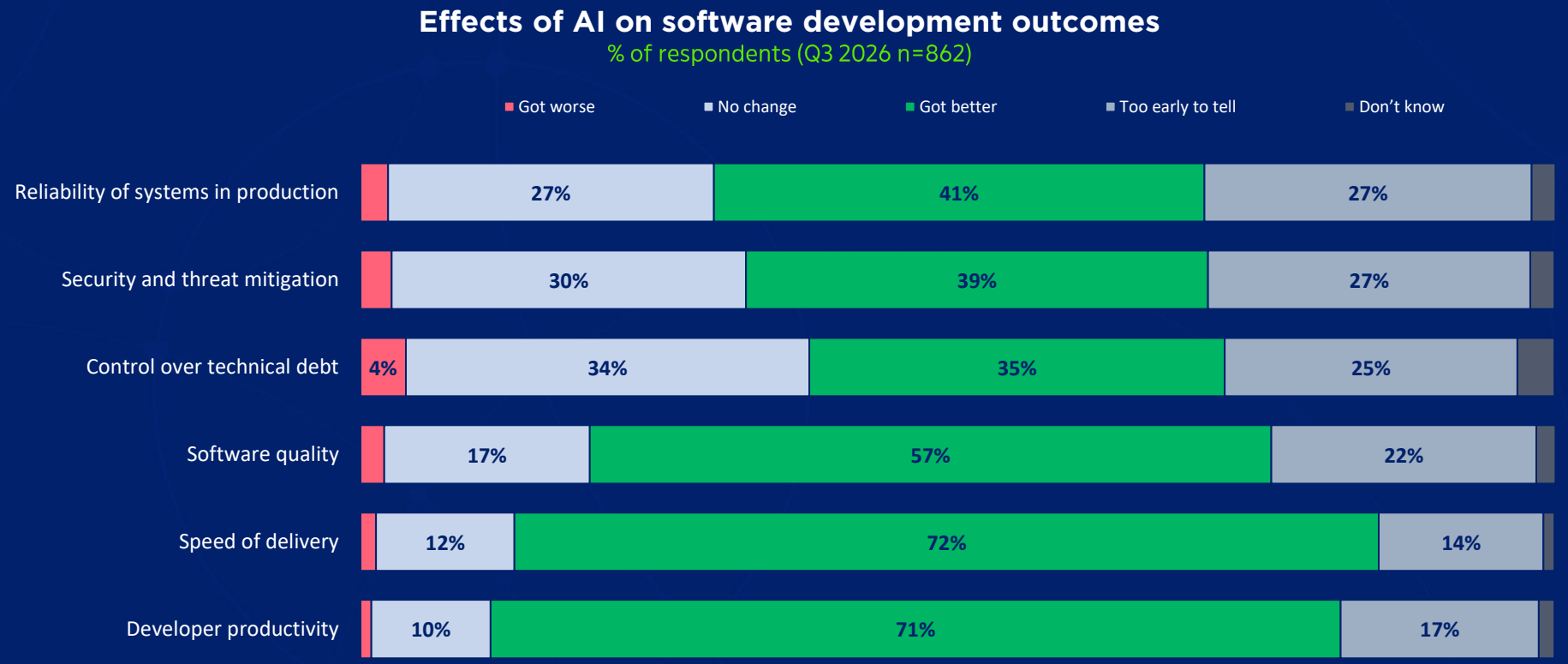
Engineering outcomes of AI on software development

Respondents report that AI has improved delivery and almost never made anything worse. 72% say delivery speed has improved since their organisation started using AI, and 71% say the same about developer productivity. 57% say software quality got better. No outcome is reported as having gotten worse by more than 4% of respondents.

The engineering outcomes that matter over time show less agreement. Reliability of systems in production (41%), security and threat mitigation (39%), and control over technical debt (35%) are reported as improved by fewer than half. For each, at least a quarter say it is too early to tell. Indeed, improvements in task completion or delivery speed may be easier for developers to observe in their daily work, while changes in software reliability, security, and maintainability may take longer to assess. These are self-reported perceptions from respondents involved in AI engineering.

Control over technical debt stands out as the outcome least frequently reported as improved, at 35%. A similar proportion (34%) report no change, while 24% say it is too early to assess the impact. This suggests that productivity gains from AI-assisted development have not translated consistently into perceived improvements in software maintainability. As AI tools accelerate code production, ensuring that the resulting software remains understandable, maintainable, and manageable over time may become an increasingly important consideration for engineering teams.

AI's perceived benefits are strongest for delivery speed and productivity, while improvements in technical debt management are less widely reported



Question wording: Since your organisation has started using AI, what effects have you seen on each of the following?

Return on AI

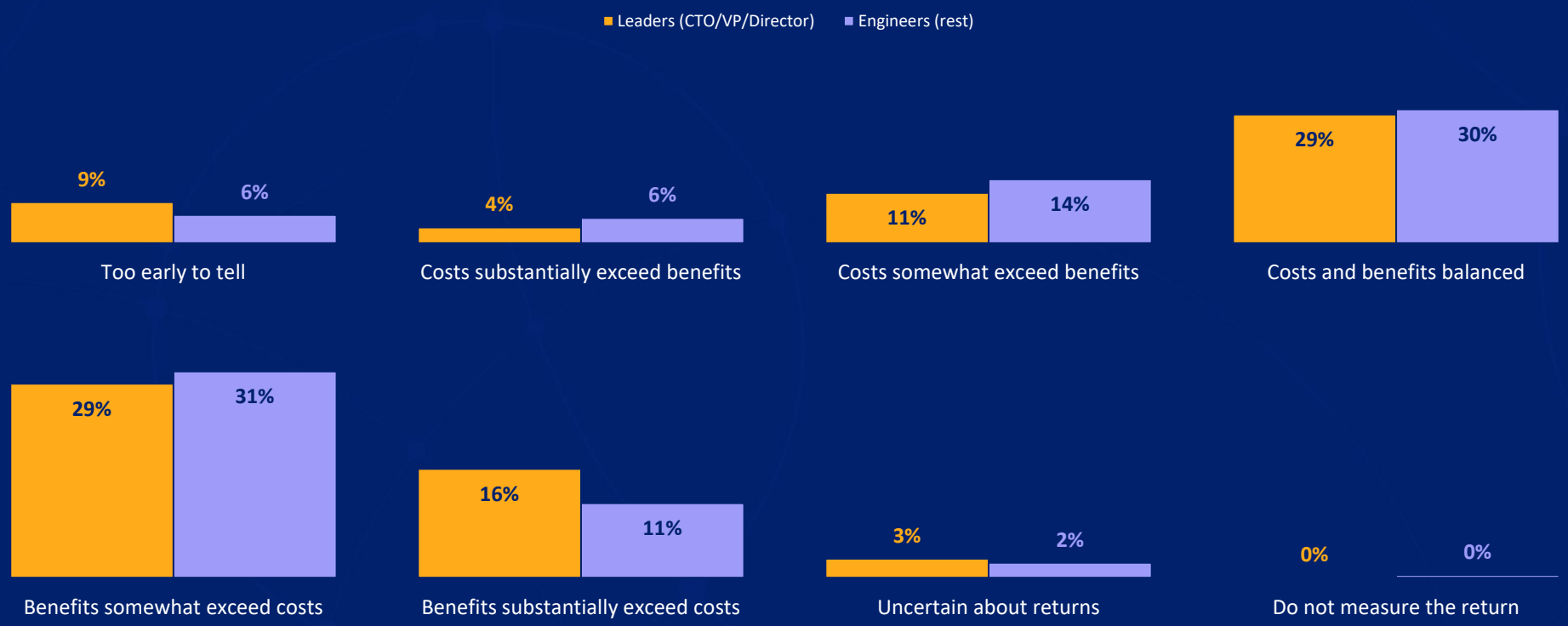
Respondents are more than twice as likely to report that AI's benefits outweigh its costs as the reverse. 42% say benefits exceed costs: 30% somewhat and 12% substantially. A further 30% say costs and benefits are broadly balanced, while 19% say costs exceed benefits.

Leaders are slightly more positive than engineers. 16% of leaders (CTOs, VPs, and directors) say benefits substantially exceed costs, against 11% of engineers. Engineers are slightly more likely to say costs exceed benefits, whether somewhat (14% against 11%) or substantially (6% against 4%). That suggests a gap between how leaders view AI's return from the top down and how engineers experience it day to day. Given the relatively small differences and the smaller leader sample, these comparisons should be treated as directional.

More respondents say AI's benefits exceed its costs than the reverse, and leaders are slightly more positive than engineers

Perceived return on AI, leaders against engineers

% of respondents (Q3 2026 n=862, leaders n=143, engineers n=719)

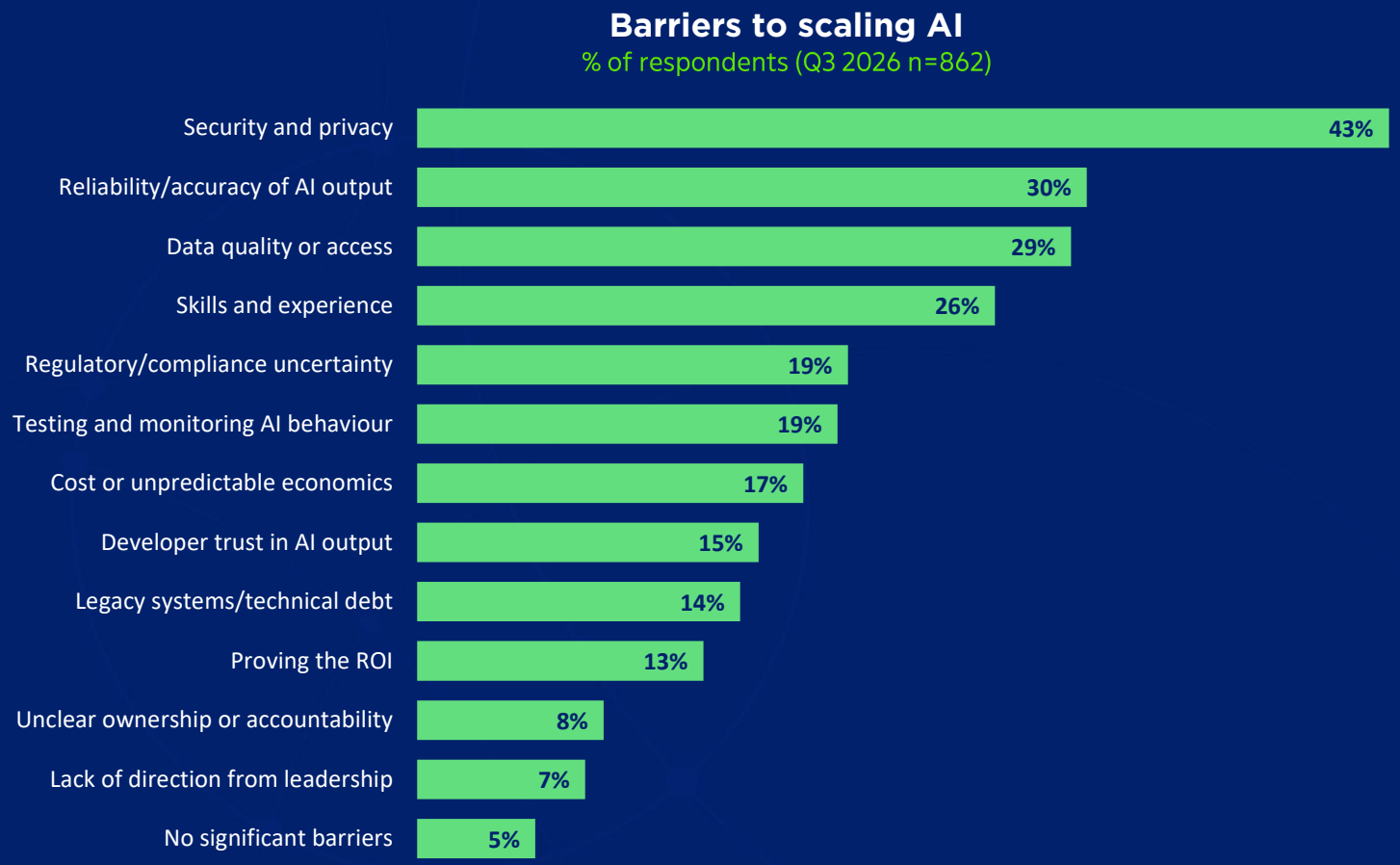


Question wording: Overall, how would you rate the value that AI brings to your organisation, relative to the cost and effort required to use it? | Which best describes your current role?

Cost considerations are less frequently prioritised as barriers to scaling AI than security, reliability, and data-related challenges. When asked to select up to three barriers, 43% of respondents chose security and privacy. Reliability or accuracy of AI output (30%) and data quality or access (29%) follow, then skills and experience (26%).

By comparison, cost or unpredictable economics (17%) and proving ROI (13%) are less frequently identified among the three biggest barriers. These findings suggest that respondents currently place greater emphasis on overcoming technical and operational constraints than on demonstrating AI's financial value. A third of organisations (34%) had AI costs significantly exceed expectations, yet few see cost as a barrier to scaling. Organisations appear to absorb cost overruns because they judge the return positive. What holds them back is whether AI can be trusted with their data and their output.

Security, reliability, and data-related challenges outweigh financial concerns as perceived barriers to scaling AI



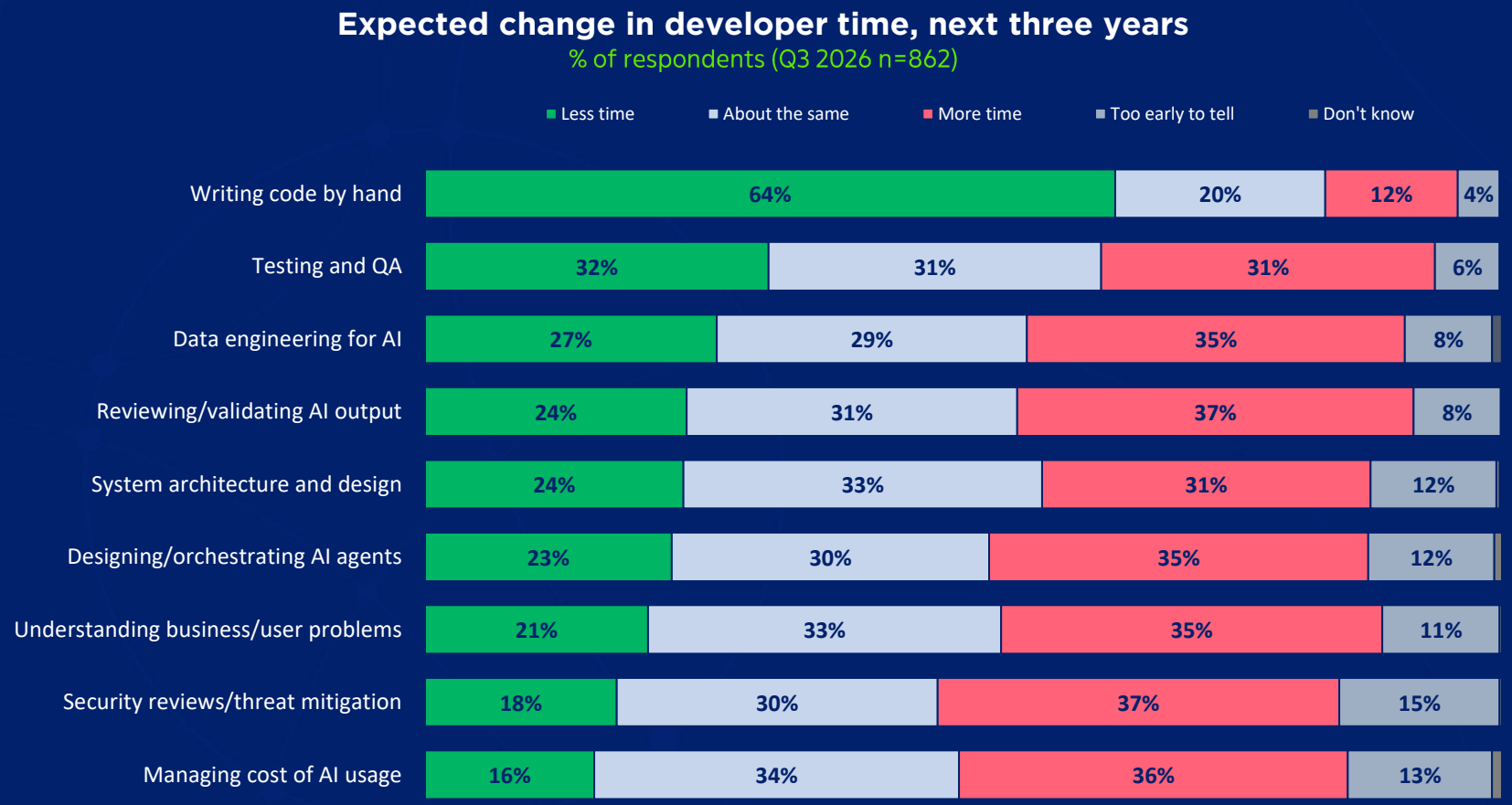
Question wording: What are the three biggest barriers to scaling AI safely and effectively in your organisation?

The developer role over the next three years

AI is making teams faster but not necessarily making software easier to maintain. Respondents agree on one change to the developer role: less time writing code by hand. 64% expect developers to spend less time on it over the next three years. No other task reaches that level of agreement. For every other task, expectations split between less time, about the same, and more time, with no answer above 37%. The tasks most often expected to take more time are security reviews and threat mitigation (37% more time, 30% about the same, 16% less time), reviewing and validating AI output (37%, 31%, 24%), and managing the cost of AI usage (36%, 34%, 16%), although they also have a high share of respondents expecting to require less time. Testing and quality assurance splits almost evenly across the three answers (32%, 31%, and 31%).

Respondents know what AI will take away from developers. They still don't agree on what it will add. The tasks expected to grow most are the ones that check and control AI, which matches the gaps identified in chapters two and three.

Two in three respondents expect developers to spend less time writing code by hand, but there is no consensus on which tasks will take its place



Question wording: Compared with today, how do you expect software developers to spend their time over the next three years?

CONCLUSIONS AND RECOMMENDATIONS

05

Progress® | /DATA

AI engineering in 2026 is concentrated in the builder stage. More than half of respondents (56%) report practices consistent with repeatable AI development, but only 16% reach the operator stage associated with systematic production engineering. The gap is widest in governance, the least advanced dimension of the AI Engineering Maturity Index. This suggests that many organisations have progressed further in adopting AI than in establishing the practices needed to manage it consistently at scale.

The practices most strongly associated with higher AI engineering maturity are structural. AI gateways, internal data pipelines, and production safeguards show some of the largest lift in maturity scores. Skills programmes and paid AI tools have spread to organisations at every stage, so they no longer distinguish the leaders. For engineering leaders, this highlights the importance of moving beyond individual use cases towards shared infrastructure, reusable integrations, and consistent operational practices as AI adoption expands.

Strengthening production governance should be a priority. Several safeguards remain relatively uncommon, including the ability to disable AI capabilities without a release, filtering outputs against defined rules, tracing model versions, logging requests and responses, documenting escalation procedures, and assigning clear ownership. These mechanisms address different aspects of operational control, from preventing undesirable behaviour to detecting, investigating, and responding to failures. Organisations moving AI capabilities into production should ensure that appropriate safeguards and accountability mechanisms are established alongside deployment.

Two areas deserve particular attention. First, AI capabilities in business processes and workflows are less likely to have reached scalable production, and advanced evaluation is less consistently accompanied by automated recovery mechanisms. As organisations expand AI into business-critical workflows, they should consider not only how capabilities are evaluated before deployment but also how their behaviour will be monitored and how failures can be managed. Second, increasing agent autonomy calls for governance approaches that combine human oversight with technical safeguards. Logging, access restrictions, sandboxing, and mechanisms to pause or revoke agent activity can complement human approval as agents take on more complex tasks.

Finally, scaling AI does not remove incidents, and it does not yet settle what developers will do next. Organisations moving AI capabilities into production should evaluate AI's value beyond productivity gains: 72% of respondents report improved delivery speed and 71% report improved developer productivity, but only 35% report improved control over technical debt. Mature organisations report more incidents, and respondents expect less hand-written code but don't agree on what replaces it. Technical debt is the outcome AI has improved least, and it may be where the next stage of AI engineering maturity is decided.

Taken together, these findings suggest that the next challenge for AI engineering is not simply accelerating development or expanding adoption but ensuring that the resulting systems can be operated reliably, securely, and sustainably. For engineering leaders, success will increasingly depend on balancing AI-driven productivity gains with the architecture, governance, skills and engineering practices needed to maintain software quality over time.

METHODOLOGY

This report presents data from a custom survey SlashData ran for Progress Software in Q3 2026. The survey collected 862 qualified responses.

Respondents were screened on five criteria. Each respondent is a professional employee involved in software development and is based outside India. Each works for an organisation that develops software, whether for customers, for its own use, or on behalf of clients. Each has at least three years of professional experience in software development. Each is involved in at least one AI engineering task beyond using AI tools for coding or day-to-day work and holds a technology role.

THE AI ENGINEERING MATURITY INDEX IS BUILT FROM 16 SURVEY QUESTIONS GROUPED INTO FIVE DIMENSIONS:



Adoption and Integration covers:

- For each area, how far has your organisation got in building or integrating AI capabilities?', asked for ‘in software we build as internal tools’, ‘in software we build for customers or clients’, ‘in our business processes and workflows’
- Which of these activities does your team typically do with help from AI coding tools?



Governance and Control covers:

- For each area, which of the following AI model(s) is used to power the AI capabilities your organisation built or integrated?
- Which of the following production safeguards apply to AI capabilities in your organisation?
- How does your organisation track and manage AI costs?
- Which of the following controls, if any, apply to AI agents in your organisation?



Evaluation and Monitoring covers:

- For each area, before release, which best describes the evaluation process typically applied to AI capabilities your organisation builds or integrates?
- For each area, after release, which best describes the monitoring process typically applied to AI capabilities your organisation builds or integrates?
- What checks apply to AI-generated code before deployment?



People and Enablement covers:

- How does your organisation support developers in building AI skills?
- Compared with today, how do you expect software developers to spend their time over the next three years?
- Since your organisation has started using AI, what effects have you seen on each of the following?
- On which of these topics does your organisation provide formal guidance or policies?



Data and Architecture covers:

- What criteria guide model choice in your organisation?
- How does your organisation prepare internal data for use in AI capabilities?
- Which architectural changes has your organisation made to support AI capabilities?

Can I share data from this report?

1. License Grant

This report is licensed under the Creative Commons Attribution-NoDerivatives Licence 4.0 (International). Put simply, subject to the terms and conditions of this license, you are free to:

Share — You can reproduce the report or incorporate parts of the report into one or more documents or publications, for commercial and non-commercial purposes.

Under the following conditions:

Attribution — You must give appropriate credit to SlashData™ and to Progress Software as sponsors of this report, and indicate if changes were made. In that case, you may do so in any reasonable manner, but not in any way that suggests that SlashData™ endorses you or your use.

NoDerivatives — you cannot remix or transform the content of the report. You may not distribute modified content.

2. Limitation of Liability

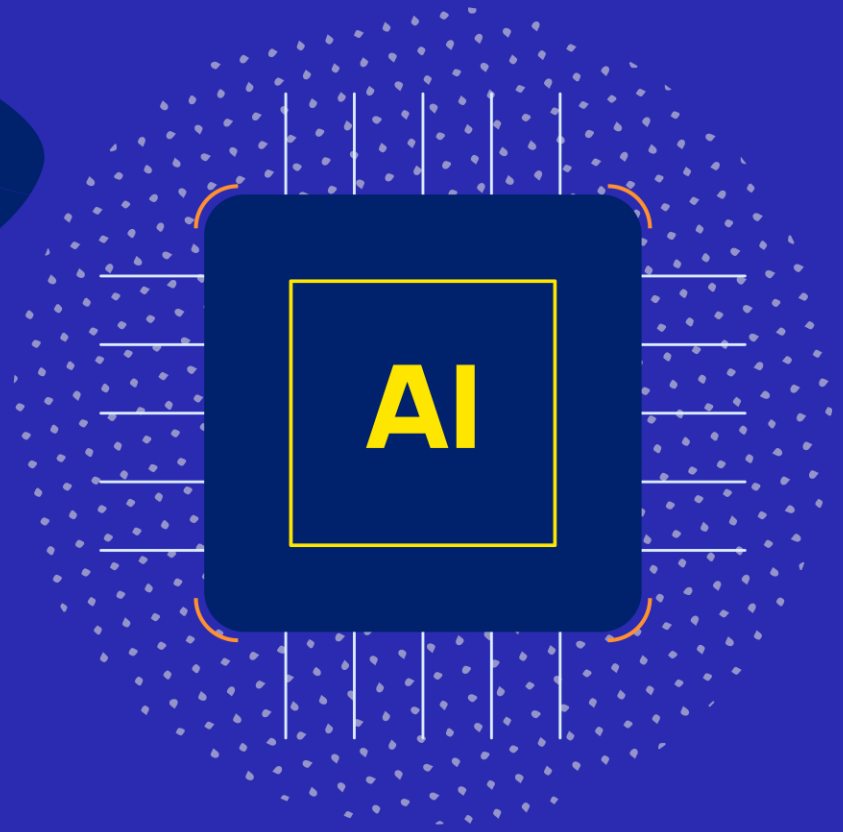
SlashData™, believes the statements contained in this publication to be based upon information that we consider reliable, but we do not represent that it is accurate or complete and it should not be relied upon as such. Opinions expressed are current opinions as of the date appearing in this publication only and the information, including the opinions contained herein, are subject to change without notice. Use of this publication by any third party for whatever purpose should not and does not absolve such third party from using due diligence in verifying the publication's contents. SlashData™ disclaims all implied warranties, including, without limitation, warranties of merchantability or fitness for a particular purpose.

SlashData™, its affiliates, and representatives shall have no liability for any direct, incidental, special, or consequential damages or lost profits, if any, suffered by any third party as a result of decisions made, or not made, or actions taken, or not taken, based on this publication.

The analyst of the developer economy | formerly known as VisionMobile

SlashData © Copyright 2026 | Some rights reserved

Progress | /DATA



Progress® | DATA