

Accelerating Software Delivery with AI

What independent evidence says about where acceleration comes from, why most organizations don't manage to accelerate, and the five things top teams do to accelerate.

Seven pages of summarized evidence for executives, board members, investors, and the advisors who brief them, with a complete list of 124 sources in the appendix. Key findings are synthesized into one actionable sentence:

To accelerate delivery, invest in your people's AI skills and enable them to reimagine your processes with AI.

The tools are necessary, but the return comes from people and process. Four findings carry that argument, and each has a section behind it.

1 **THE FACT** Only one team in twenty gets to 2x acceleration with AI.

- Across 22,000 organizations, the median team's delivery throughput rose 4% with AI; the top 5% of teams roughly doubled (CircleCI).
- Individual engineers using AI in their own way are 10–25% faster on real work, but the organization typically gets 0–10%: coding is only 25–35% of idea-to-launch, and the gains that do appear are absorbed downstream by review queues, larger changes, rework, and incidents.
- Engineers also overestimate their own gain — self-assessed productivity correlates with measured output at just 0.17 (Stanford).

→ Sections 1 and 2

2 **THE CAUSE** People and process impact the result 2x-4x more than technology.

- Four independent studies agree: a clear AI strategy lifts measurable impact by about 25 points against about 5 for better tools (BCG); organizational factors explain 67% of AI impact against 32% for individual factors (Microsoft); 73% of AI high performers redesigned workflows against 25% of others (McKinsey); technology is about 20% of an initiative's value (PwC).
- And it does not fix itself: 81% of organizations provide AI tools, 46% provide training; 84% of developers use AI, 29% trust its output; every tool generation makes unmanaged use faster and messier — bugs per developer +9% → +54%, median review time +91% → +441% (Faros).

→ Section 3

3 **THE MOVES** The teams that accelerate do five common things, and none of them is buying another tool.

- Leadership decides at the top — strategy, investment, a named initiative, a baseline and target, and where saved time goes.
- People are coached hands-on, on real work, and the practice is captured so it lives across the team, not in two people.
- Intent is made explicit — context, specs, standards, and rules files become the new control surface.
- Review and verification are re-engineered for AI volume — risk-tiered review, tests with every change, a human in control.
- Workflows are redesigned before the tools scale—the whole lifecycle, with product management and QA inside it.

→ Section 4

4 **THE PAYOFF** With those five in place, the same teams deliver 2x+.

- Organizations that redesign how work flows reach +25% to +30% productivity and faster time-to-market (Bain, McKinsey).
- The top 5% of teams roughly double (CircleCI).
- An enterprise mandate paired with standardized tooling and re-engineered review reached 2.09x throughput with merge and revert rates stable (Stanford/CMU, 802 developers).

→ Section 5

How to read the report. Section 1 gives the ranges and the confidence behind them; Section 5 puts the payoff in one table (Table 3). Section 2 explains why individual gains do not reach the organization. Section 3 gives the cause and the five barriers, one number each. Section 4 outlines the five things the teams that accelerate do, with evidence for each. Section 5 is the five questions and the payoff. The appendix lists all 124 sources with links.

Prepared by Mladen Malkovich · Partnered Solutions LLC · aisprints.ai · Evidence window 2022 – September 2026, weighted toward controlled trials and large-scale telemetry published since September 2025. This digest may be shared in full; please retain this notice and the source appendix.

THE FACT

1. The fact: only 1 out of 20 teams gets to 2x

Nearly every engineer now uses AI. Very few organizations use it the same way, to a standard, with results that show up in delivery.

Individual engineers using AI in their own way get measurably faster, and the organization gets almost nothing: the evidence reviewed across 124 sources puts the organization-level return on ad-hoc use at roughly +10%, says the value above it is where most of the return sits, and says what separates the teams that accelerate from the rest is people and process, not tooling.

Acceleration does not come from a better tool license; it comes from growing the structured practice. Every source that documents that step names the same enablers: coaching, shared specs and standards, small batches, AI-adapted review, a clear organizational AI stance, and leadership that decides where saved time goes.

How confident is this?

- High for individuals (multiple controlled trials):
 - GitHub/Accenture +8.7% PRs
 - Google +21%
 - Cui et al. +26% across 4,867 developers;
 - METR's expert maintainers on million-line repos –19% in 2025 and –18%/–4% in the 2026 rerun, both with confidence intervals spanning zero.
- High for the team-level paradox (five or more independent large datasets agree).
- Medium for the structured-practice step:
 - Consistent across consultancies, telemetry vendors, and one large quasi-experimental study (the Stanford/CMU 2x mandate, 802 developers, 196k PRs, 2.09x throughput with stable merge and revert rates)
 - No controlled trial randomizes “structured vs ad-hoc”
 - Self-reported figures run about 20 points optimistic.

The following two figures visualize data reported by the most relevant sources, both for individual engineers and for teams and organizations.

Individual engineers: AI acceleration reported by study, 2022–2026

AI + Ad-hoc Use clusters at +10–25% on real work and has drifted upward; AI + Structured Practice reaches +50% to ~2x

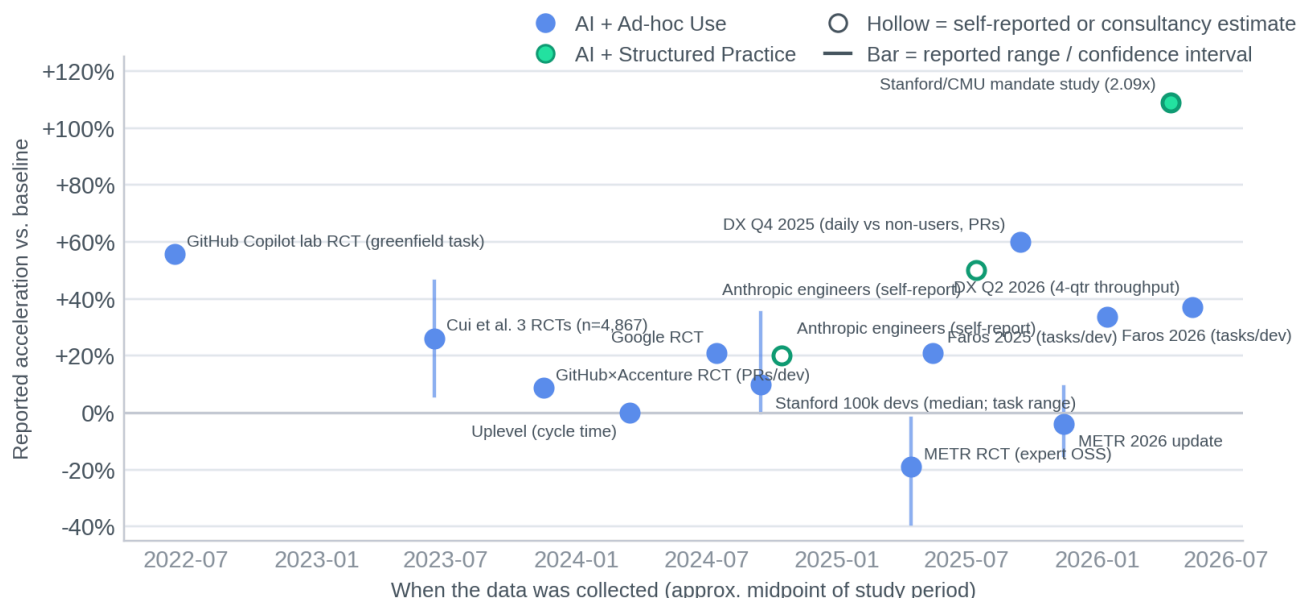


Figure 1. Individual engineers: AI acceleration reported by study, 2022–2026. Each point is one study’s headline figure at the approximate midpoint of its data collection. Blue = AI + Ad-hoc Use; green = AI + Structured Practice; hollow markers are self-reported or consultancy estimates; bars show reported ranges or confidence intervals.

Teams and organizations: AI acceleration reported by study, 2024–2026

AI + Ad-hoc Use sits near 0–15% at the organization level; AI + Structured Practice and top-decile teams reach +25% to ~2x

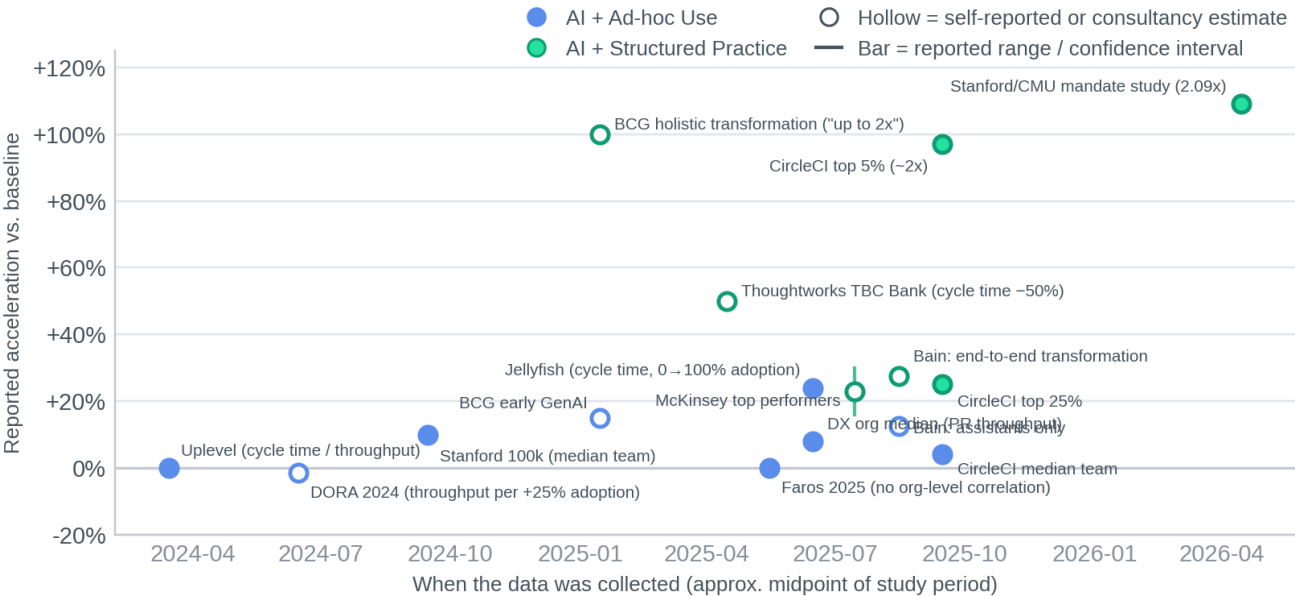


Figure 2. Teams and organizations: AI acceleration reported by study, 2024–2026. Ad hoc adoption sits near 0–15% at the organization level; structured practice and the best-performing teams reach +25% to ~2x. Metrics differ across studies; the pattern, not any single point, is the finding.

The trend is widening the gap.

- Individual uplift roughly doubled between late 2024 and mid-2026, driven mainly by the shift from autocomplete to agentic tools plus accumulated user skill rather than model releases:
 - Stanford/CMU study finds its gains “not attributable to specific model generations”, and METR flags a learning curve of roughly 50 hours before gains appear.
- Over the same period, the quality cost grew faster than the speed gain, and the team-level median barely moved while the top 5% pulled away to ~2x (Table 1).
- Each tool generation makes ad-hoc use faster and messier, so structured practice becomes more valuable with every release, not less.

Metric	Earlier reading	Later reading(s)	Source
Self-reported time saved per developer	~2 h/week (Q4 2024)	3.6 h average (Q4 2025) → 6+ h heavy users (Q2 2026)	DX
Anthropic engineers, self-reported productivity	+20% (2024)	+50% (2025); 14% report >100%	Anthropic
Share of merged code that is AI-generated	22% (Q4 2025)	24% (Q1 2026) → 52.7% (Q2 2026)	DX
Bugs per developer	+9% (2025)	+54% (2026)	Faros
PR review time	+91% (2025)	Median +441%; time to first review +157%	Faros
Incidents per PR / code churn	—	+243% (~3x) / +861%	Faros 2026
Review queue for agentic PRs	—	~5x longer pickup (16+ h vs ~200 min); merge rate 32.7% vs 84.5% for human PRs	LinearB 2026
Team throughput, 22k organizations	—	Aggregate +59% y/y; median team +4%; top 25% +25%; top 10% +47%; top 5% ~2x	CircleCI 2026

Table 1. The trend in numbers: per-engineer speed keeps climbing with each tool generation; effective output (speed net of rework, review load, and incidents) climbs much less, and the gains concentrate in a small share of teams.

THE FACT, CONTINUED

2. Why individual gains do not reach the organization

- Coding is only 25–35% of idea-to-launch time (Bain), and developers spend about 14% of their time coding (DX), so even a 30% coding speedup bounds out near +5–10% end to end.
- Ad-hoc AI use then adds cost downstream: pull requests about 2.5x larger (LinearB) and +154% (Faros); reviewer pickup roughly five times slower and median review time +441%; churn +861% and rework 2.6x (Faros, Stanford); bugs +9% → +54% and incidents per PR ~3x; and 31% more PRs merged without any review.
- Several large studies find no organization-level correlation between AI adoption and delivery at all (DORA 2024: –1.5% throughput and –7.2% stability per 25 points of adoption; Faros 2025: no significant correlation; Uplevel: no change in cycle time, +41% bugs).
- Self-perception inflates the picture: METR’s developers believed they were 20% faster while measured 19% slower, and Stanford finds only a 0.17 correlation between self-assessed and measured productivity.
- Gains are also highly task-dependent: 0–10% on complex brownfield work against 30–35% on simple greenfield tasks (Stanford).
- “How do your engineers feel about AI” is the wrong survey; “what is delivery throughput and rework against baseline” is the right one.

THE CAUSE

3. The cause: people and process, and the five barriers that hold most organizations back

Research over the last twelve months is unusually consistent on one point: the organizational system around the tools, not the tools, decides whether value is realized.

Four independent studies, using four different methods, put the people-and-process lever at two to four times the technology lever (Figure 4).

Five barriers explain why so few organizations pull that lever, and why ad hoc use persists on its own.

1. Organizations buy tools and skip the rest.

- Across 204 organizations in 37 countries, 81% provide AI tool access, 46% provide training, 41% have governance policies, and 21% have dedicated AI experts (Giray et al.).
- Only 12% of employees say they receive sufficient AI training (EY), and 90% of those trained found it insufficient (HFS).
- McKinsey’s top accelerators use embedded coaching about 30% more often than the rest, and 57% of top performers invest in hands-on workshops and coaching, compared with 20% of bottom performers; the rest rely on the license.

2. Trust is low, and verification is the new bottleneck.

- 84% of developers use or plan to use AI, but 29% trust its accuracy, down 11 points in a year (Stack Overflow)
- 96% do not fully trust AI code to be correct, yet only 48% always check it before committing (Sonar).
- The check has moved downstream: median time in review is up 441%, 31% more PRs merge with no review (Faros).
- In open source, 61% of 33,596 AI-generated pull requests received no recorded review (Duma et al.).
- Without an AI-adapted review design, the team’s senior engineers become the constraint (Figure 3).

Near-universal use, minority trust
% of developers in each survey

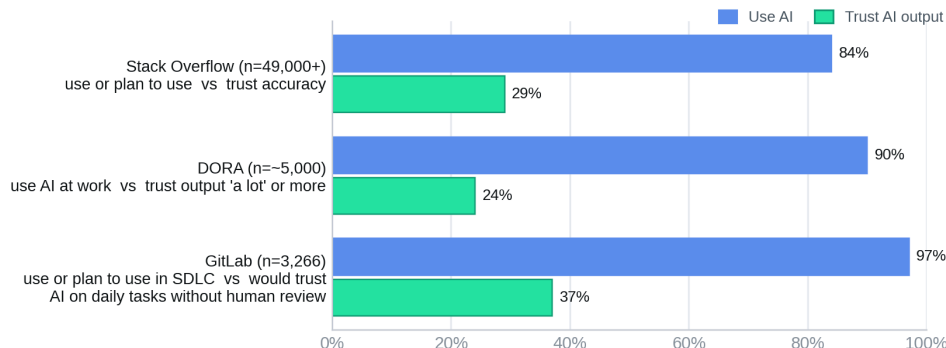


Figure 3. Near-universal use, minority trust: share of developers using AI versus trusting its output, three surveys (Stack Overflow n=49,000+; DORA n=5,000; GitLab n=3,266).

3. Saved time leaks because nobody decided where it goes.

- 42% of frontline users save a workday or more each week, but 66% receive no guidance on using the time (BCG).
- Only 13% of workers are rewarded for reinventing work with AI, and 45% say it feels safer to focus on current goals than to redesign how work is done (Microsoft). The incentive to stay ad hoc is built in.

The people-and-process lever moves AI impact far more than the technology lever

Four independent studies; each row compares the human/process factor with the tooling or individual factor

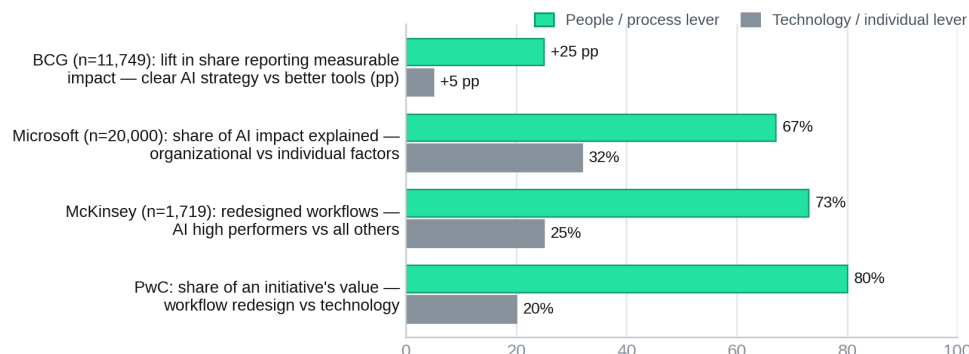


Figure 4. The people-and-process lever moves AI impact far more than the technology lever. A clear AI strategy lifts the share of employees reporting measurable impact by about 25 points against about 5 for better tools (BCG, n=11,749); organizational factors explain 67% of AI impact against 32% for individual factors (Microsoft, n=20,000); 73% of McKinsey's AI high performers have fundamentally redesigned workflows against 25% of others; PwC puts technology at about 20% of an initiative's value.

4. Leaders and practitioners see different realities.

- 81% of CEOs are highly confident in AI-driven systems, compared with 56% of QA and DevOps professionals.
- 44% of C-level respondents believe the organization can govern AI agents, compared with 23% of QA staff (Tricentis).
- 89% of executives say AI increased speed, but 6% can show organization-wide ROI (Atlassian); leaders set targets and measurement regimes from the rosier picture.
- 54% of developers fear AI-derived metrics will be used for individual evaluation (Harness), which is itself a reason to under-report problems.

5. The change is landing on people, and they are pushing back.

- 41% of employees report increased cognitive load and 47% now spend more time managing and directing AI than doing the work (BCG).
- Employee resistance to AI rose from 5% to 20% in a single quarter (KPMG).
- Only 33% of frontline staff say leadership communicates clearly about AI (BCG).
- Prosci's benchmark puts the cost of getting this wrong at 88% project success with excellent change management against 13% with poor, and 73% with active executive sponsorship against 29% without.

Taken together, these barriers make ad hoc use a stable equilibrium rather than a stage teams pass through.

Getting out of it requires the same things the acceleration evidence associates with accelerating teams: a leadership-level decision, a redesign of how the team works, and hands-on enablement that changes how people use the tools. Section 4 sets those out as five moves.

THE MOVES

4. The moves: the teams that accelerate do five common things.

The sources that document teams moving from AI + Ad-hoc Use into AI + Structured Practice name overlapping enablers. Grouped, they come down to five common things (Table 2).

The move	What it means in practice	Evidence
1 LEADERSHIP Leadership decides at the top	A stated AI strategy and a clear organizational AI stance (DORA); an investment and a named initiative; a baseline before anything starts and a target after; an explicit decision on where saved time goes (quality, security, modernization, scope); incentives that reward redesigned work; outcome, quality and trust metrics rather than adoption metrics, with developers involved in defining them.	66% of employees get no guidance on using saved time (BCG); only 13% are rewarded for reinventing work (Microsoft); ~80% of McKinsey's top performers link AI goals to performance reviews; 73% vs 29% success with active executive sponsorship (Prosci).
2 PEOPLE People are coached hands-on, on real work	Coaching and mentoring on the committed backlog rather than passive training, until each person has the competence to direct AI and the confidence to use it on their own work; calibrated trust rather than blind trust or blanket distrust; AI used for inquiry rather than delegation so skills form rather than atrophy; cognitive load and approval fatigue managed so speed does not cost competence; people able to explain what they did and why, and to vouch for the result.	57% vs 20% invest in coaching (McKinsey); 90% found generic training insufficient (HFS); inquiry-style use preserved learning while delegation did not (Anthropic RCT); only 10% of organizations combine strong enablement with high adoption (Jellyfish); 41% report higher cognitive load (BCG).
3 PRACTICE Intent is made explicit	Spec-driven development and codified project knowledge; context, specs, standards and rules files become the new control surface, version-controlled alongside the code and written down so the practice lives across the team, not in two people; requirements precise enough to survive a literal-minded executor, since "ambiguity travels faster in an AI-enabled PDLC and can do more damage" (McKinsey); codebase coherence and context quality so agents can work safely; debt triage and "AI slop" prevention as churn rises.	Documentation and rules files worth about 3x the incremental throughput; coherent codebases gained ~4x while fragmented ones went slightly negative (Jellyfish); churn +861% and rework 2.6x where practice did not change (Faros, Stanford); spec-driven governance (Farrag).
4 QUALITY Review and verification are re-engineered for AI volume	AI-adapted, risk-tiered review; tests with every change; quality at the authoring stage; automated review and "autonomous validation" as volume rises, with a human in control; platform-level guardrails and golden paths; named accountability for AI quality; logs of agent actions and explicit human-AI boundaries (HFS; Deloitte).	Elite organizations run AI code review on 57% of PRs against 8% at fair tier (LinearB); quality held in the 2x mandate only where review was re-engineered (Stanford/CMU); 85% say agentic AI succeeds best under a platform-engineering approach (GitLab); 73% have hardly any standardized templates or golden paths, and heavy users report more incidents and longer MTTR (Harness).
5 WORKFLOWS Workflows are redesigned before the tools are scaled	Workflows and roles redesigned across the whole lifecycle, with product management and QA inside the practice rather than downstream (Bain); small batches, strong version control, and task-suitability judgment (DORA's seven capabilities; Stanford task-type data); shared specs and acceptance criteria across roles.	93% of McKinsey's top accelerators redesigned workflows first; organizations that redesign processes before deploying AI are more than twice as likely to report gains above 20% (McKinsey); DORA 2025 finds seven capabilities that amplify AI's effect and flip its sign on team performance.

Table 2. The moves that enable some teams to accelerate 2x+ with the same tools and people.

Product and QA are inside the practice, not outside.

- In quality engineering, 89% of organizations are piloting or deploying GenAI, but only 15% have scaled it enterprise-wide (World Quality Report); 60% report deploying untested code, 30% of them because AI-generated volume overwhelms testing capacity (Tricentis); testers who actively use AI are 17% less anxious about their profession than those who do not (PractiTest); and Forrester describes the tester's shift "from scripting tests to setting quality goals and supervising testing agents."
- In requirements and product management, organizations with over 50% AI adoption report their largest savings upstream, in requirements, user stories, and ideation (Gartner); human-AI collaboration accounts for 54% of AI use in requirements engineering and full automation for 5% (Rani et al.); and Forrester names the product manager's critical skill as "the ability to provide clear intent, context, and constraints to AI peers."
- A structured practice that changes only the developers changes only a third of the lifecycle.

THE PAYOFF

5. The payoff, and five questions to ask your technology leaders and teams

With those five moves in place, the same teams deliver 2x+.

Table below (Table 3) summarizes the payoff for the teams that made 5 moves mentioned earlier. **The tool license is the same in both right-hand columns of Table 3**; what changes between AI + Ad-hoc Use and AI + Structured Practice is the five moves above.

	AI + Ad-hoc Use	AI + Structured Practice
Individual engineers	+10% to +25% faster on real work (controlled trials cluster at +8–26%; median telemetry 10–15%).	+25% to +50% effective output; ~2x throughput in the best-documented program: roughly double the ad-hoc gain.
Teams and organizations	~0% to +10% at the organization level; individual gains are absorbed by review queues, larger changes, rework and incidents; delivery stability typically gets worse.	+25% to +30% increased productivity and faster time-to-market for organizations that redesign how work flows (Bain, McKinsey); +47% for the top 10% of teams and ~2x for the top 5% (CircleCI); 2.09x in the Stanford/CMU 2026 mandate study.

Table 3. The summary of payback for individuals and teams that pursue 5 moves.

That's why the practice is worth more than any single tool: the practice is the asset, and the vendor changes every few months.

“We use AI” no longer distinguishes anything. These questions tell you in a single conversation whether a team is accelerating or still at the median.

- 1. Show me delivery throughput and change-failure rate for the last 12 months. What changed in how the team works, and when?** If the answer is a tool name, expect 0–10% at the company level (DORA, Faros, CircleCI).
- 2. Who verifies AI-generated code, and how long does a change wait for review?** Agentic PRs wait roughly five times longer to be picked up (LinearB), and 31% more often merge without review (Faros). If one or two senior engineers are the gate, that is the scaling constraint.
- 3. Are product management and QA inside the re-designed AI workflow, or downstream of it?** Coding is 25–35% of the lifecycle (Bain); 60% of organizations are deploying untested code (Tricentis). If only developers changed, only a third of the value changed.
- 4. Is the practice written down, actually practiced, and continually improved, or does it live in a few people?** Key-person risk in a new form: what happens if the two best AI users leave? 81% of organizations provide tools; 46% provide training (Giray et al.).
- 5. What are we spending on AI tokens and licenses, and can we tie it to measured results against a baseline, and to a decision about where saved time goes?** Roughly half of CIOs cannot quantify impact (BCG); self-assessment correlates with measured output at only 0.17 (Stanford); 66% of employees get no guidance on using saved time (BCG 2026).

THE EVIDENCE

Evidence and how to read it

Sources were gathered in September 2026 from analyst firms, strategy consultancies, engineering-intelligence vendors, research groups, and AI-native companies, and screened for a stated method, a stated sample and a quantified result.

Sources: evidence is weighted in the following tiers.

1. Randomized or controlled trials (GitHub/Accenture, Cui et al., Google, METR, Anthropic).
2. Large-scale telemetry (Stanford/CMU, Faros, DX, CircleCI, LinearB, Jellyfish).
3. Structured surveys and consultancy analyses (DORA, Bain, McKinsey, BCG, Microsoft, Stack Overflow, Sonar, Tricentis).
4. Vendor and company self-reports.

Headline ranges lean on the first two tiers; the third and fourth guide the barriers, and the five moves.

The time trend uses data-collection dates, not publication dates.

The acceleration evidence uses a 2022–2026 window because the trend needs pre-2025 baselines; the human-and-process evidence is restricted to sources published between September 2025 and September 2026.

Caveats: most team-level data comes from vendors analyzing their own customer bases, and consultancies sell transformation; no controlled trial randomizes “structured vs ad-hoc”, so that estimate is observational and quasi-experimental; self-reported hours saved run high and should be treated as upper bounds; PR counts inflate as PRs shrink or agents open them, and “share of code AI-written” is an input, not an outcome; quality trends worsened with agentic tools in 2026, so any acceleration claim should be paired with a quality metric. DORA 2025 (“foundations amplify”) and Faros 2026 (“foundations don’t protect”) disagree; the Stanford/CMU study sits between: quality held only where review was re-engineered alongside the mandate.

THE SOURCES

Appendix. Complete list of sources reviewed (124)

All 124 sources were reviewed for this digest and accessed 12–13 September 2026; each had to state a method, a sample, and a quantified result to be included. Entries are grouped by organization type. Sources quoted in the text, tables, or figures are set in black; sources set in grey were reviewed and inform the ranges, the barriers, and the five moves without being quoted directly. † marks pre-September-2025 acceleration studies kept as baselines for the time trend in Figures 1–2 and Table 1.

Analyst firms (Gartner, Forrester, IDC-class)

1. **Gartner (Philip Walsh)** — Don't Limit AI in Software Engineering to Coding (9 Dec 2025). [gartner.com](https://www.gartner.com)
2. **Gartner** — Predicts 2026: AI Potential and Risks Emerge in Software Engineering Technologies (paywalled; figures quoted via ArmorCode reprint, 22 Dec 2025) (3 Dec 2025). [gartner.com](https://www.gartner.com) · armorcode.com
3. **Gartner press release (n=360 IT application leaders)** — Gartner Survey Finds Just 15% of IT Application Leaders Are Considering, Piloting, or Deploying Fully Autonomous AI Agents (30 Sep 2025). [gartner.com](https://www.gartner.com)
4. **Gartner press release** — Gartner Unveils Top Predictions for IT Organizations and Users in 2026 and Beyond (21 Oct 2025). [gartner.com](https://www.gartner.com)
5. **Gartner** — Enterprise AI Coding Agents: 2026 Market Guide & Trends (Apr 2026). [gartner.com](https://www.gartner.com)
6. **Gartner press release** — Gartner Says the Market for Enterprise AI Coding Agents Is Entering a New Phase of Expansion and Competitive Realignment (20 May 2026). [gartner.com](https://www.gartner.com)
7. **Gartner press release** — Gartner Predicts AI Coding Costs Will Surpass Average Developer's Salary by 2028 as Token Consumption Surges (24 Jun 2026). [gartner.com](https://www.gartner.com)
8. **Gartner press release** — Gartner Predicts 60% of Organizations Will Adopt Smaller Software Engineering Teams by 2029 (7 Jul 2026). [gartner.com](https://www.gartner.com)
9. **Forrester (Diego Lo Giudice)** — Predictions 2026: Software Development Goes From Jamming To A Full Orchestra (4 Nov 2025). [forrester.com](https://www.forrester.com)
10. **Forrester (Diego Lo Giudice et al.)** — AI Is Rewriting Software Work: What It Means For Your Team (11 Nov 2025). [forrester.com](https://www.forrester.com)
11. **Forrester (Diego Lo Giudice)** — Agentic Software Development Takes The Lead: From Code Assistants To Orchestrated SDLC Agents (c. Mar 2026). [forrester.com](https://www.forrester.com)
12. **Forrester (Diego Lo Giudice)** — The Autonomous Testing Platform Wave Q4 2025 Is Out / Beyond The Wave: What Customers Really Think About Autonomous Testing Platforms (9 Jan 2026 / 2026). [forrester.com](https://www.forrester.com) · [forrester.com](https://www.forrester.com)

Consulting and advisory firms

13. **McKinsey (Relyea, Lunawat, Palaniappan, Harrysson, Linderman, Dixit)** — Beyond the copilot: Scaling the agentic product development life cycle (21 Aug 2026). [mckinsey.com](https://www.mckinsey.com)
14. **McKinsey (Moon, Walsh, Di Leo, Thelwall)** — Rewiring software delivery for the agentic era (28 May 2026). [mckinsey.com](https://www.mckinsey.com)
15. **McKinsey (Relyea, Harrysson)** — The AI revolution in software development (excerpt from Rewired, 2nd ed.) (1 Apr 2026). [mckinsey.com](https://www.mckinsey.com)
16. **McKinsey / QuantumBlack** — The State of AI: Global Survey 2026 (n=1,719, 97 countries) (25 Aug 2026). [mckinsey.com](https://www.mckinsey.com)
17. **BCG** — AI at Work 2026: Strategy Matters More Than Tools (n=11,749 workers, 14 markets) (3 Jun 2026). [bcg.com](https://www.bcg.com) · [web-assets.bcg.com](https://www.web-assets.bcg.com)
18. **BCG** — AI Fuels Field Service Transformation. Change Management Provides the Spark (restates BCG's 10-20-70 rule) (9 Jan 2026). [bcg.com](https://www.bcg.com)
19. **Bain & Company** — From Pilots to Payoff: Generative AI in Software Development (Technology Report 2025) (23 Sep 2025). [bain.com](https://www.bain.com)
20. **Deloitte Insights** — Tech Trends 2026 (incl. 'The agentic reality check') (10 Dec 2025). [deloitte.com](https://www.deloitte.com) · [deloitte.com](https://www.deloitte.com)
21. **Deloitte** — State of AI in the Enterprise 2026 (n=3,235 leaders, 24 countries; fielded Aug-Sep 2025) (21 Jan 2026). [deloitte.com](https://www.deloitte.com)
22. **Deloitte Insights** — 2026 Global Software Industry Outlook (12 Feb 2026). [deloitte.com](https://www.deloitte.com)
23. **PwC US** — 2026 AI Business Predictions (2 Dec 2025). [pwc.com](https://www.pwc.com)
24. **KPMG US** — AI Quarterly Pulse Survey, Q2 2026 (n=204 US C-suite/leaders, \$1B+ firms) (24 Jun 2026). [kpmg.com](https://www.kpmg.com)
25. **EY US** — EY survey: autonomous AI adoption surges at tech companies as oversight falls behind (n=500 US tech leaders) (4 Mar 2026). [ey.com](https://www.ey.com)
26. **EY Global** — Work Reimagined Survey 2025: companies missing up to 40% of AI productivity gains (n=15,000 employees + 1,500 employers, 29 countries) (10 Nov 2025). [ey.com](https://www.ey.com)
27. **Capgemini, Sogeti, OpenText** — World Quality Report 2025-26 (17th ed.; 2,000+ senior executives, 22 countries) (13 Nov 2025). [capgemini.com](https://www.capgemini.com) · [capgemini.com](https://www.capgemini.com)
28. **Info-Tech Research Group** — AI Adoption & Impact Study: AI in Software Development, June 2026 (n=578 application/engineering/product leaders) (30 Jun 2026 (insights) / 20 Jul 2026 (press release)). [infotech.com](https://www.infotech.com) · [infotech.com](https://www.infotech.com)
29. **HFS Research (Hansa Iyengar)** — Reliable AI at scale requires trust engineering, not legacy QE (9 Sep 2026). [hfsresearch.com](https://www.hfsresearch.com)
30. **HFS Research (Jhunjhunwala, Chaturvedi)** — Scale agentic AI without a learning loop and you scale failure (21 Aug 2026). [hfsresearch.com](https://www.hfsresearch.com)
31. **Thoughtworks** — Technology Radar Vol. 33 - 'Compacency with AI-generated code' (Hold) (5 Nov 2025). [thoughtworks.com](https://www.thoughtworks.com) · [thoughtworks.com](https://www.thoughtworks.com)

32. **Thoughtworks** — Technology Radar Vol. 34 - cognitive debt, 'putting coding agents on a leash' (15 Apr 2026). [thoughtworks.com](https://www.thoughtworks.com) · [thoughtworks.com](https://www.thoughtworks.com)
33. **Thoughtworks (Richard Gall)** — Macro trends in the tech industry, April 2026 (15 Apr 2026). [thoughtworks.com](https://www.thoughtworks.com)
34. **Thoughtworks (Liu Shangqi)** — Spec-driven development: Unpacking one of 2025's key new AI-assisted engineering practices (4 Dec 2025). [thoughtworks.com](https://www.thoughtworks.com)
35. **Prosci** — 5 Strategic Decisions for Building Organizational Change Capability in 2026 (5 Nov 2025). [prosci.com](https://www.prosci.com)
36. **Carnegie Mellon Software Engineering Institute (Ozkaya et al.)** — From Hype to Adoption: Guiding Organizations in Their AI Journey / Managing the Complexities of AI Adoption (10 Nov 2025 / 1 Jun 2026). sei.cmu.edu · sei.cmu.edu

Research institutes, academic and peer-reviewed studies

37. **Google Cloud DORA** — State of AI-assisted Software Development 2025 (~5,000 respondents + 100 hours qualitative) (23 Sep 2025). dora.dev · services.google.com
38. **Google Cloud DORA; coverage by InfoQ** — The ROI of AI-assisted Software Development (2026) (11 May 2026). cloud.google.com · infoq.com
39. **METR** — We are Changing our Developer Productivity Experiment Design (24 Feb 2026). metr.org
40. **METR** — Measuring the Self-Reported Impact of Early-2026 AI on Technical Worker Productivity (n=349) (11 May 2026). metr.org
41. **Anthropic (Shen, Tamkin); arXiv 2601.20245** — How AI assistance impacts the formation of coding skills (RCT, n=52) (29 Jan 2026). anthropic.com · arxiv.org
42. **Anthropic** — How AI is transforming work at Anthropic (132 surveyed, 53 interviews, 200k Claude Code transcripts) (2 Dec 2025). anthropic.com
43. **Anthropic** — Anthropic Economic Index: Learning curves (March 2026 report) (24 Mar 2026). anthropic.com
44. **Anthropic** — Anthropic Economic Index: Cadences (June 2026 report; ~9,700 linked survey respondents) (26 Jun 2026). anthropic.com
45. **Cui, Demirer, Jaffe, Musolf, Peng, Salz - Management Science (Articles in Advance)** — The Effects of Generative AI on High-Skilled Work: Evidence from Three Field Experiments with Software Developers (n=4,867) (27 Feb 2026). pubsonline.informs.org
46. **Vella, Blincoe - arXiv 2605.23135** — The Impact of AI Coding Assistants on Software Engineering: A Longitudinal Study (22 May 2026). arxiv.org
47. **Chen, He, Williams, Valentino, Talwalkar - arXiv 2602.03593** — Beyond the Commit: Developer Perspectives on Productivity with AI Coding Assistants (n=2,989 at BNY) (3 Feb 2026). arxiv.org
48. **Giray, Demirors, Kalinowski, Mendez - arXiv 2512.23327** — An Empirical Study of Generative AI Adoption in Software Engineering (n=204, 37 countries) (29 Dec 2025 (v2 1 Apr 2026)). arxiv.org
49. **Duma et al. - arXiv 2605.02273** — These Aren't the Reviews You're Looking For: How Humans Review AI-Generated Pull Requests (EASE 2026; 33,596 AI PRs) (4 May 2026). arxiv.org
50. **Huang et al. - arXiv 2601.21276** — More Code, Less Reuse: Investigating Code Quality and Reviewer Sentiment towards AI-generated Pull Requests (MSR 2026) (29 Jan 2026). arxiv.org
51. **Haider, Zimmermann - arXiv 2601.19287** — Understanding Dominant Themes in Reviewing Agentic AI-authored Code (MSR 2026; 19,450 review comments) (27 Jan 2026). arxiv.org
52. **Agarwal, Miller, Kastner, Vasilescu (CMU) - arXiv 2607.07980** — 3100 Opinions on Code Review in an AI World: Building Causal Theory from Practitioner Discourse (8 Jul 2026). arxiv.org
53. **Bappy et al. - arXiv 2605.23130** — From Preventive to Reactive: How AI Coding Assistants Transform Developers' Security Awareness (v2 10 Jun 2026). arxiv.org
54. **Huang, Reyna, Lerner, Xia, Hempel - arXiv 2512.14012** — Professional Software Developers Don't Vibe, They Control: AI Agent Use for Coding in 2025 (16 Dec 2025 (v2 18 Aug 2026)). arxiv.org
55. **Pimenova, Fakhoury, Bird, Storey, Endres (Microsoft Research / Uvic) - arXiv 2509.12491** — Good Vibrations? A Qualitative Study of Co-Creation, Communication, Flow, and Trust in Vibe Coding (Sep 2025 (v2 29 Jun 2026)). arxiv.org
56. **Houck, Choudhuri, Kalliamvakou, Zimmermann - ACM Queue** — The AI-Native Developer (survey of 1,300+ developers, 22 interviews) (May 2026). queue.acm.org
57. **Santana, Magalhaes, de Souza Santos - arXiv 2510.17164** — Software Testing with Large Language Models: An Interview Study with Practitioners (n=15) (20 Oct 2025). arxiv.org
58. **Yoshimoto et al. - arXiv 2603.13724** — Testing with AI Agents: An Empirical Study of Test Generation Frequency, Quality, and Coverage (14 Mar 2026). arxiv.org

59. You, Dong, Wang, Li, Peng - arXiv 2607.24000 — Industrial Practice of LLM-Based Test Case Carving and Assertion Generation (ISSTA 2026 experience paper) (27 Jul 2026). [arxiv.org](#)
 60. Murali Rani, Berntsson Svensson, Feldt (Chalmers) - arXiv 2511.01324 — AI for Requirements Engineering: Industry Adoption and Practitioner Perspectives (n=55) (3 Nov 2025). [arxiv.org](#)
 61. Hess, Vogelsang, Franch, Herrmann, Kopczynska, Rachmann - Springer — Opportunities and Limitations of GenAI in RE: Viewpoints from Practice (REFSQ 2026) (25 Mar 2026). [link.springer.com](#)
 62. IREB Requirements Engineering Magazine (Chetan Arora, Monash) — Ethics of Using LLMs in Requirements Engineering (18 Nov 2025). [re-magazine.ireb.org](#)
 63. Jin, Jin et al. - arXiv 2602.18306 — ReqElicitGym: An Evaluation Environment for Interview Competence in Conversational Requirements Elicitation (20 Feb 2026). [arxiv.org](#)
 64. Bhati (Northeastern) - arXiv 2604.26275 — Agentic AI in the Software Development Lifecycle: Architecture, Empirical Evidence, and the Reshaping of Software Engineering (29 Apr 2026). [arxiv.org](#)
 65. Harvard Business Review (Niederhoffer et al.) — AI-Generated 'Workshop' Is Destroying Productivity (BetterUp Labs + Stanford Social Media Lab; n=1,150) (22 Sep 2025). [hbr.org](#)
 66. Harvard Business Review (Ranganathan, Ye - UC Berkeley Haas) — AI Doesn't Reduce Work - It Intensifies It (9 Feb 2026). [hbr.org](#)
 67. Harvard Business Review (Srinivasan, Wei) — To Thrive in the AI Era, Companies Need Agent Managers (12 Feb 2026). [hbr.org](#)
 68. Brynjolfsson, Chandar, Chen - Stanford Digital Economy Lab — Canaries in the Coal Mine? Six Facts about the Recent Employment Effects of AI (Aug 2026 update) (12 Aug 2026). [digitaleconomy.stanford.edu](#)
 69. Microsoft (Jared Spataro) — 2026 Work Trend Index: How Frontier Firms are rebuilding the operating model for the age of AI (20,000 workers, 10 countries) (5 May 2026). [blogs.microsoft.com](#)
 70. He, Agarwal, Denisov-Blanch et al. (Stanford / CMU) - arXiv 2607.01904 — AI Writes Faster Than Humans Can Review: A Longitudinal Study of an Enterprise 2x Mandate (802 developers, 196,212 PRs; difference-in-differences) (Jul 2026). [arxiv.org](#)
- ### Industry surveys, platform telemetry and practitioner sources
71. Stack Overflow (Eira May) — Mind the gap: Closing the AI trust gap for developers (analysis of 2025 Developer Survey, n=49,000+) (18 Feb 2026). [stackoverflow.blog](#)
 72. GitHub — Octoverse 2025: A new developer joins GitHub every second as AI leads TypeScript to #1 (28 Oct 2025). [github.blog](#)
 73. JetBrains — The State of Developer Ecosystem 2025 (n=24,534) (15 Oct 2025). [blog.jetbrains.com](#) · [devecosystem-2025.jetbrains.com](#)
 74. JetBrains — AI Coding Agents: Adoption Trends (Developer Ecosystem Survey 2026 early release; n=15,000+) (19 Aug 2026). [blog.jetbrains.com](#)
 75. GitLab — Global DevSecOps Survey 2025: The Intelligent Software Development Era (n=3,266; The Harris Poll) (10 Nov 2025). [about.gitlab.com](#) · [businesswire.com](#)
 76. Sonar — State of Code Developer Survey Report 2026: the verification gap (n=1,100+) (8 Jan 2026). [sonarsource.com](#)
 77. Atlassian Teamwork Lab — State of Teams 2026: The AI efficiency paradox (12,000+ knowledge workers, 170+ Fortune 1000 executives) (15 May 2026). [atlassian.com](#)
 78. Atlassian — The AI-native SDLC is paying off: 19% more PRs and 2-3 hours saved per developer per week (31 May 2026). [atlassian.com](#)
 79. DX — AI Impact Report Q1 2026 (400+ engineering organizations) (15 Apr 2026). [getdx.com](#)
 80. DX — The State of AI Impact in Engineering: Q2 2026 (500+ engineering organizations) (22 Jul 2026). [getdx.com](#)
 81. Faros AI — AI Engineering Report 2026: The Acceleration Whiplash (~22,000 developers, 4,000+ teams) (12 Apr 2026). [faros.ai](#) · [pages.faros.ai](#)
 82. LinearB — 2026 Software Engineering Benchmarks Report / 8 million pull requests reveal... (8.1M PRs, 4,800 teams) (4 May 2026). [linearb.io](#) · [linearb.io](#)
 83. GitClear (GitKraken) — The Maintainability Gap: 2026 AI Code Quality Research (623M code changes) (Jan 2026). [gitclear.com](#)
 84. Veracode — 2026 GenAI Code Security Report (100+ LLMs) (28 Jul 2026). [veracode.com](#)
 85. Jellyfish — 2026 State of Engineering Management Report (n=600+ engineering leaders) (12 May 2026). [jellyfish.co](#)
 86. Jellyfish Research — The State of AI in Software Engineering (H1 2026; 37M PRs, 134,911 developers) (Jun 2026). [jellyfish.co](#)
 87. Harness — State of DevOps Modernization 2026 (n=700; Coleman Parkes) (11 Mar 2026). [prnewswire.com](#)
 88. Harness — The State of Engineering Excellence 2026 (n=700; Sapio Research) (13 May 2026). [prnewswire.com](#)
 89. LeadDev — The Engineering Leadership Report 2026 (n=600 engineering leaders) (1 Sep 2026). [leaddev.com](#)
 90. Lenny's Newsletter (Noam Segal) — How tech workers actually feel about AI in 2026 (annual AI sentiment survey) (12 Jul 2026). [lennysnewsletter.com](#)
 91. Tricentis — 2026 Quality Transformation Report (n=2,501; Censuswide) (3 Jun 2026). [tricentis.com](#)
 92. Ranorex / Sembli — Software Quality Pulse Report, first edition (~4,000 respondents) (19 May 2026). [ranorex.com](#)
 93. Perforce (Panterra) — 2026 State of DevOps Report (n=820; 54% C-level) (24 Feb 2026). [prnewswire.com](#)
 94. BrowserStack — State of AI in Software Testing 2026 (n=250+ testing leaders) (10 Feb 2026). [prnewswire.com](#)
 95. Leapwork / InfoQ — Leapwork research on AI in testing (n=300+; via InfoQ) (17 Feb 2026). [infoq.com](#)
 96. Applause — 2025 State of Digital Quality Report (n=2,100+) (17 Sep 2025). [businesswire.com](#)
 97. PractiTest with Tea-time with Testers — The 2026 State of Testing Report (13th edition) (Q1 2026 (exact release date not shown)). [practitest.com](#)
 98. ISTQB — ISTQB announces minor update to Certified Tester - Testing with Generative AI (CT-GenAI v1.1); CT-AI v2.0 syllabus (27 Apr 2026 / May 2026). [istqb.org](#) · [istqb.org](#)
 99. General Assembly — Product Managers Are All-in on AI, but Skills Gaps and Shadow AI Pose Risks (n=117 PMs) (6 Nov 2025). [businesswire.com](#)
 100. ProductPlan — The State of Product Management Report 2026 (n=250 product leaders) (2026 (month not shown)). [productplan.com](#)
 101. Silicon Valley Product Group (Marty Cagan) — The AI Productivity Paradox (23 Jul 2026). [svpg.com](#)
 102. IIBA (Jared Gorai) — Top 6 Business Analysis Trends to Monitor in 2026 (8 Jan 2026). [iiba.org](#)
 103. Pluralsight — Pluralsight Unveils 2026 Tech Forecast: AI Hype Deflates, Skills Take Center Stage (11 Nov 2025). [pluralsight.com](#)
 104. martinoflower.com (Birgitta Böckeler, Thoughtworks) — Harness engineering for coding agent users (2 Apr 2026). [martinoflower.com](#)
 105. Addy Osmani (Google) — Agentic Engineering / Own the Outer Loop (4 Feb 2026 / 15 Jul 2026). [addyosmani.com](#) · [addyosmani.com](#)
 106. Simon Willison — Writing about Agentic Engineering Patterns (23 Feb 2026). [simonwillison.net](#)
- ### Acceleration studies and time-trend baselines, 2022–2026 († = published before September 2025)
107. GitHub — [Research: quantifying GitHub Copilot's impact on developer productivity and happiness](#) (2022)†
 108. GitHub — [Quantifying GitHub Copilot's impact in the enterprise with Accenture](#) (2024)†
 109. Paradis et al. (Google) — [How much does AI impact development speed? An enterprise-based randomized controlled trial](#) (2024)†
 110. McKinsey — [Unleashing developer productivity with generative AI](#) (2023)†
 111. METR — [Measuring the Impact of Early-2025 AI on Experienced Open-Source Developer Productivity](#) (Jul 2025)†; [arXiv](#)
 112. Denisov-Blanch (Stanford) — [Will AI Replace Software Engineers? 100k-developer study slides](#) (2025)†
 113. Faros AI — [The AI Productivity Paradox](#) (Jul 2025)†
 114. DX — [AI-assisted engineering: Q4 2025 impact report](#) (Jan 2026)
 115. Uplevel — [Can generative AI improve developer productivity?](#) (2024)†
 116. Google DORA — [2024 Accelerate State of DevOps†; RedMonk analysis](#)
 117. McKinsey — [Unlocking the value of AI in software development](#) (Nov 2025)
 118. McKinsey — [AI-native development: rethinking software from the ground up](#) (Jul 2026)
 119. BCG — [Executive Perspectives: AI-Enabled Engineering Excellence](#) (Apr 2025)†; [BCG X: AI-assisted coding and generative engineering](#)
 120. CircleCI — [2026 State of Software Delivery](#)
 121. Atlassian — [State of Developer Experience 2025](#) (Jul 2025)†
 122. Thoughtworks — [The business benefits of AI-first software engineering](#) (2025)†
 123. Farrag — [The Productivity-Reliability Paradox: Specification-Driven Governance](#) (May 2026)
 124. AWS — [Amazon Q Developer just reached a \\$260 million milestone](#) (2024)†
- Third-party statistics and short quotations are cited from their original publishers and remain the property of those publishers; no third-party charts or slides are reproduced. All trademarks belong to their respective owners. This digest may be shared in full; please retain this notice and the source appendix.