A&M | PRIVATE EQUITY
PERFORMANCE IMPROVEMENT

AI in Software Development

Why AI Tooling Adoption Alone Will Not Move the Needle

Executive Summary

Artificial intelligence (AI) has achieved near-universal penetration inside engineering organizations. Adoption rates of AI tooling among developers are reported to be **up to 90%** at some firms, and executive commentary from companies including Spotify and Uber describes a fundamental shift in how software is being written.[1,2,7,9] However, adoption and tooling alone don't equal value creation. Rigorous, independent measurement tells a more complicated story: time saved has plateaued at roughly **four hours per engineer per week** with no sustained upward trend.[2] In one closely watched study, developers perceived speed gains of 20% to 24% while objective measurement of the same cohort revealed a **19% overall slowdown**.[6]

The gap between perception and proven results is apparent and it is not a technology problem. It is a **measurement and process problem**, with variations in reporting and some firms unable to properly define how they measure gains. To help bridge this gap, A&M recommends that software development teams adopt a common set of practices to demonstrate value creation using AI in their software development lifecycle (SDLC), including:

- Establishing a credible baseline of software development metrics before deploying tools
- Quantifying key performance indicators (KPIs) for overall product and software development
- Treating the adoption of AI in the SDLC as a process redesign as inseparable from AI tooling adoption

This paper describes what those practices look like, why they matter, and how organizations can realize real value with AI-assisted engineering.

The Signal Behind the Noise

The narrative around AI-assisted engineering has fractured into two camps: enthusiasts who describe transformational gains and skeptics who see little beyond marginal productivity improvements. Even some of the best engineering teams are citing wildly different productivity gains, and some are even admittingly making up those figures.[5] The divergence doesn't reflect a disagreement about AI's potential, but rather a disagreement over the conditions under which that potential is, or is not, realized.

What the research consensus does establish clearly is this: **high adoption rates are not, on their own, predictive of business impact**. The 2025 DORA State of AI-Assisted Software Development report, the DX AI-Assisted Engineering Q4 Impact Report, and independent research from METR collectively document a pattern in which developer sentiment and self-reported productivity diverge sharply from objective, instrumented outcome data.[5,6,7] Engineers feel faster; however, cycle times and defect rates, where measured, often do not confirm it.

The common denominator among organizations failing to close this gap is the absence of measurement infrastructure. They cannot demonstrate value not because value is absent but because they have no process in place for detecting it.

> R&D organizations treating AI as a tooling exercise, rather than a value creation initiative, consistently fail to translate adoption into improved financial results. It all starts with baselining key metrics to track improvements.

What Organizations Need To Do Differently

Across high-performing engineering organizations, three practices consistently distinguish those generating measurable financial returns from AI investment from those that are not.

1. Diagnose Before Prescribing Tooling

Most organizations select AI tooling and then ask whether it is working. They begin by assessing **SDLC maturity, current AI adoption depth, and the availability of baseline performance data** before any intervention is recommended. This diagnostic step serves two purposes. First, it reveals where AI is likely to produce the fastest and most measurable gains—which are typically in code generation, automated testing, and continuous integration and continuous delivery—versus where investment is likely to underdeliver, such as upstream architectural and strategy functions where feedback loops are slower and outputs are harder to measure. Second, it generates the baseline against which future progress can be objectively assessed.

Without a diagnostic baseline, improvement is anecdotal. With a baseline, outcomes become both measurable and auditable.

2. Build a Measurement Foundation

Best-in-class organizations instrument their development lifecycle around a consistent set of KPIs that are mapped directly to financial outcomes rather than developer satisfaction surveys or license utilization rates. Cycle time, for example, is not merely an engineering efficiency metric—it is a direct input into **revenue realization velocity and EBITDA expansion**. Under A&M's proprietary approach, the metrics that matter most fall into seven categories:

The common denominator among organizations failing to close this gap is the absence of measurement infrastructure. They cannot demonstrate value not because value is absent but because they have no process in place for detecting it.

KPI	What It Measures	Why It Links to P&L
Cycle Time	End-to-end development to production deployment	Faster cycle times accelerate revenue realization from new features. With the same cost base, organizations can realize higher output
Engineering Throughput	Completed stories per sprint per full-time equivalent, normalized for scope	Directly measures cost per feature and enables headcount redeployment without quality loss
Defect Escape Rate	Production defects as a percentage of total defects	Reduces customer support costs and incident response while protecting revenue retention
Automated Test Coverage	Percentage of software testing that is automated	Lowers manual quality assurance labor and post-release remediation, stabilizing delivery costs
Deployment Frequency	How often code reaches production	Faster, smaller releases reduce release risk and enable faster feature monetization
Technical Debt Ratio	Engineering time on maintenance versus new development	Lower debt reduces ongoing drag and improves sustainable EBITDA margin over time
Percent Roadmap Delivered On Time	Delivery predictability versus plan	Supports revenue forecast reliability and reduces contractor and fire-drill spending

The discipline to map each of these metrics to a specific financial outcome before the engagement begins is required. Organizations that make this link explicit from the outset can report AI's impact in terms of margin, cost structure, and enterprise value to boards and investors.

3. Treat Process Change as Non-Negotiable

Purchasing AI tools and distributing licenses does not change how software gets built. Redesigning workflows, team structures, and decision rights is what really impacts the development process. As a result, AI capabilities need to be deliberately embedded at each stage to achieve the most impact.

The highest-performing engineering organizations have made an explicit shift. Developers are no longer primarily writing code; they are **orchestrating AI agents**, reviewing and validating outputs, and focusing human judgment on the decisions machines cannot yet make. This transition requires deliberate change management, and engineers need structured support to adapt their working patterns. Without it, most will default to using AI as an expensive autocomplete feature, yielding marginal gains rather than long-term structural improvements.

Organizations that embed change management into AI deployment from day one consistently outperform those that treat it as an afterthought once adoption stalls.

> The highest-impact AI deployments are not distinguished by the sophistication of their tooling. They are distinguished by the discipline of their process redesign, and rigor of their measurement systems.

Connecting Engineering Performance to Enterprise Value

For PE-backed companies in particular, the strategic case for AI in engineering is not primarily about developer productivity. It is about what engineering performance does to the income statement and, ultimately, to enterprise value at exit.

This is accomplished through mechanisms such as:

- **Faster cycle time** enables faster feature monetization
- **Higher throughput at constant headcount** directly reduces research and development cost as a percentage of revenue
- **Lower defect escape rates** reduce customer support burden and protect net revenue retention
- **Improved delivery predictability** reduces working capital volatility and increases forecast confidence for board reporting

None of these outcomes are automatic consequences of AI adoption. They require the upstream work of baselining, KPI instrumentation, and process redesign described above to be realized. Organizations that do that work first, before measuring returns, are the ones generating credible results that can be reported to boards and that can attract prospective acquirers.

Measuring AI Impact and How A&M Can Help

Implementing AI tooling alone in the SDLC is not a silver bullet, but it is a genuine lever that, when applied with measurement rigor and process discipline, can produce material improvements in engineering cost structure, revenue velocity, and software quality. The firms generating those improvements are not the ones that moved fastest on dabbling with AI tools. They are the ones that asked the hard question first: **how will we know it's working?**

A&M can help answer that question with our AI in SDLC framework, which establishes a baseline, a KPI framework tied to financial outcomes, and an organizational change program running in parallel with tool deployment. This approach directly ties value creation outcomes to the AI investments being made by engineering teams.

---

1. Spotify Q4 2025 Earnings Call
2. 93% of Developers Use AI. Why Is Productivity Only 10%?
3. The Diary of a CTO, February 2026
4. Bessemer Venture Partners citing Head of Engineering at Shopify (Farhan Thawar). April 2026
5. DORA: 2025 State of AI-Assisted Software Development
6. METR: Measuring the Impact of Early-2025 AI on Experienced Open-Source Developer Productivity
7. DX: AI-Assisted Engineering Q4 2025 Impact Report
8. ICONIQ: Average Relative Increase in Productivity for Use Cases Where AI Support Is Being Deployed
9. Uber Technologies, Inc. - Uber Expands AI Data Platform to Power Next-Gen Enterprise and AI Lab Needs

Authors:

Robert Kreft
Managing Director
rkreft@alvarezandmarsal.com

Anil Kumar
Managing Director
anil.kumar@alvarezandmarsal.com

Michael Iacona
Senior Director
miacona@alvarezandmarsal.com

Azmat Mohammed
Senior Director
amohammed@alvarezandmarsal.com

Charles Ramseyer
Director
cramseyer@alvarezandmarsal.com

Harith Kesari
Director
hkesari@alvarezandmarsal.com

ABOUT ALVAREZ & MARSAL

Founded in 1983, Alvarez & Marsal is a leading global professional services firm. Renowned for its leadership, action and results, Alvarez & Marsal provides advisory, business performance improvement and turnaround management services, delivering practical solutions to address clients' unique challenges. With a worldwide network of experienced operators, world-class consultants, former regulators and industry authorities, Alvarez & Marsal helps corporates, boards, private equity firms, law firms and government agencies drive transformation, mitigate risk and unlock value at every stage of growth.

To learn more, visit: **AlvarezandMarsal.com**

Follow A&M on:

ALVAREZ & MARSAL
LEADERSHIP. ACTION. RESULTS.℠