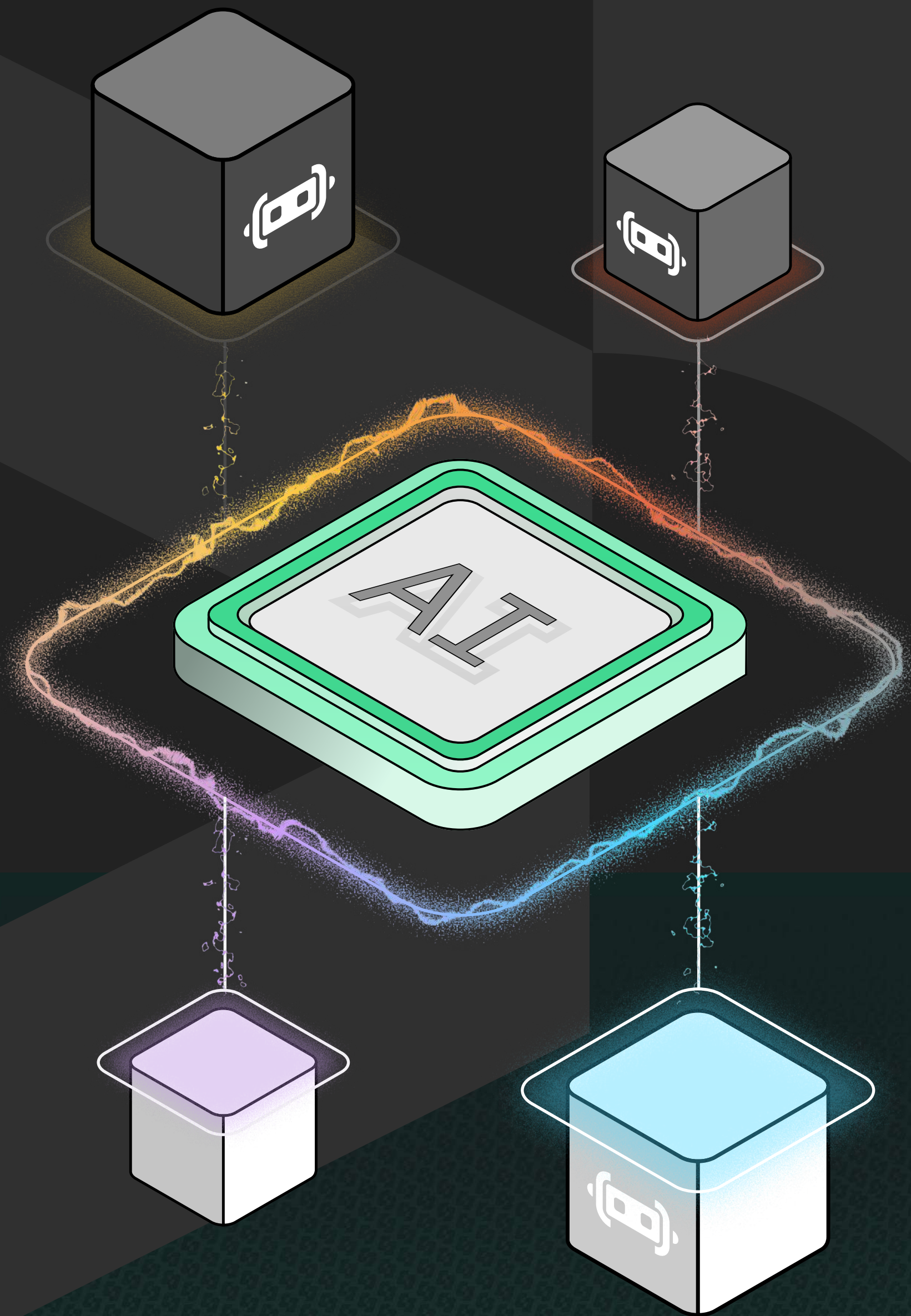


The 2025 Software Testing Vibe Check: Agentic AI Edition

How software leaders are feeling about agentic AI, and who they blame when agents go rogue





Say the phrase ‘agentic AI’ and you will elicit all sorts of reactions.

But no matter your position, the near-term future of artificial intelligence appears to be agentic: systems capable of operating without prompts to achieve specific goals and business outcomes. Agents can theoretically analyze situations, plan actions or sequences, and execute tasks based on live results and feedback.

Given the undeniable potential of this technology, many companies are investing in AI agents across various applications and industries. In software testing, AI agents could theoretically perform an increasing number of functions that currently fall to engineering and R&D departments. But what happens if fully autonomous testing becomes viable? Could it introduce as many new problems and complexities as the technology is meant to solve?

The following guide examines these issues using survey data collected from software testing executives and leaders. The questions cover anticipated benefits that company leaders and engineering team leaders expect to gain from AI agents and some of the pitfalls and limitations.

AI trends don’t exist in isolation, though. To align with these findings, we also asked about other non-AI workplace trends that may contribute to overall attitudes and behaviors (such as how companies handle accountability and decision-making).

EVP

CTO

VP

CEO

CIO

SVP

CSO

Director

CEO

CIO

CIO

VP

COO

CTO

SVP

Methodology

US-BASED SURVEY 

In June 2025, Sauce Labs partnered with Wakefield Research* to commission an online survey of:

200

U.S. Executives / C-Suites
(CTOs, CIOs, CEOs, and similar leaders)

200

U.S. Engineering Leaders
(Directors/Associate Director and VP, SVP, EVP of Engineering)

*[Note: Results of any sample are subject to sampling variation. The magnitude of the variation is measurable and is affected by the number of interviews and the level of the percentages expressing the results. For the total 400 interviews conducted in this particular study, the chances are 95 in 100 that a survey result does not vary, plus or minus, by more than 4.9 percentage points from the result that would be obtained if interviews had been conducted with all persons in the universe represented by the sample.]

START HERE:

The disconnect between leadership and software testing teams

Before diving into the full survey results, a troubling statistic sets the stage: 61% of respondents believe their top leadership lacks true understanding of software testing requirements, suggesting that critical AI adoption decisions may be made without sufficient technical insight. (The sentiment is slightly higher among engineering leaders than executives).

While the rest of the data helps determine benchmarks about where your program stands and where you may want to go, the divide between software testers, leaders, and executives is a non-starter. Teams and leaders should unite to examine and agree to improve this dynamic before making any significant decisions about agentic AI usage or changes in current testing practices.



61%

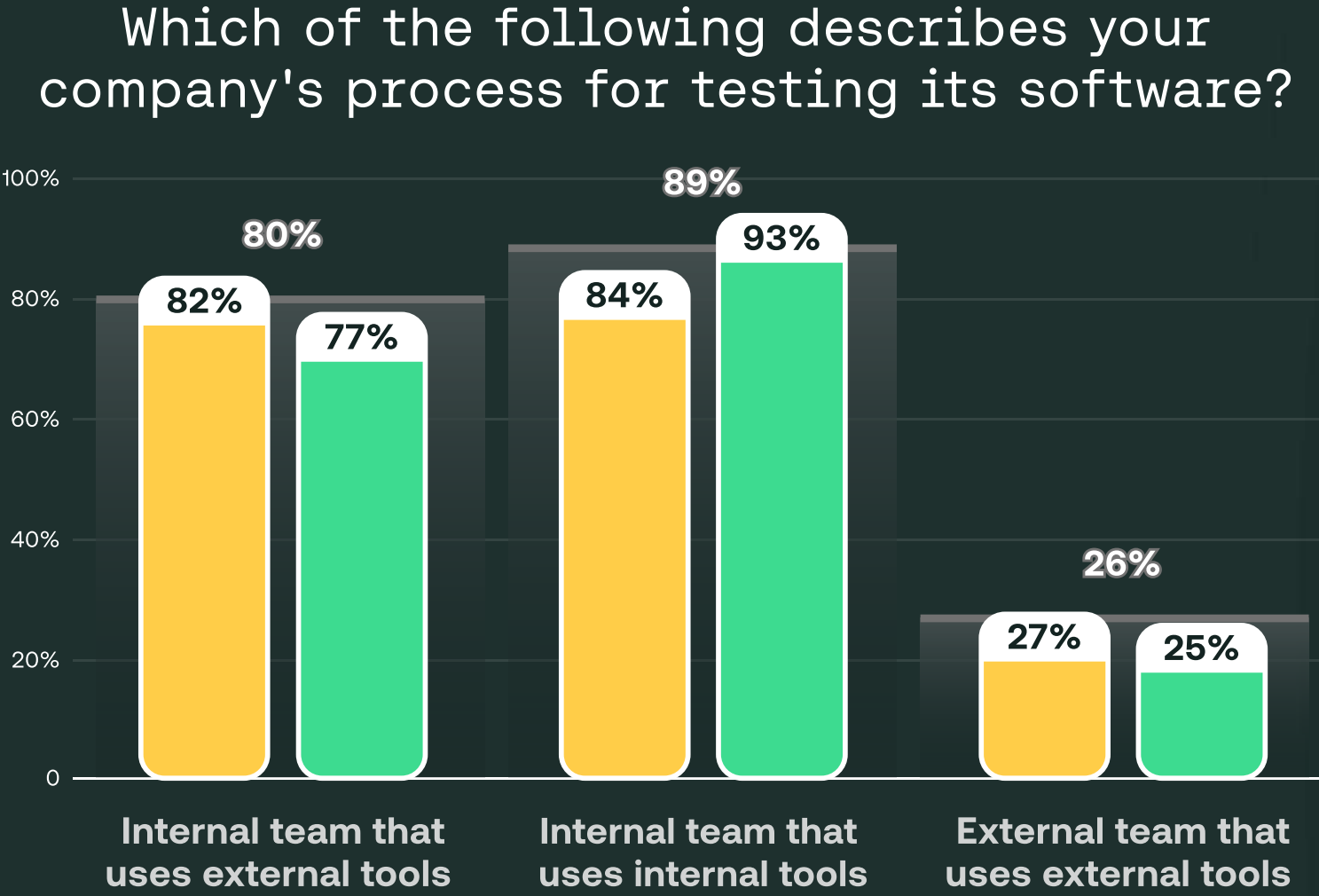
Believe their top leadership lacks true understanding of software testing requirements

65% vs. 57%

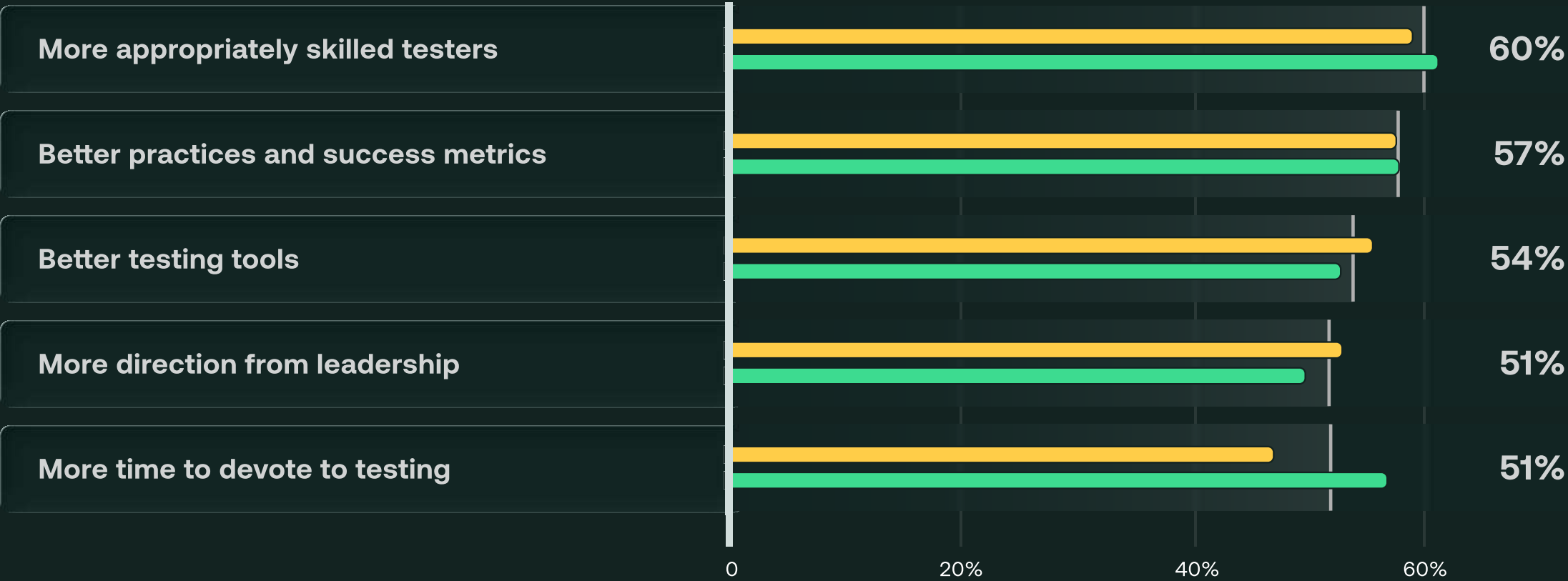
The sentiment is slightly higher among engineering leaders than executives

Most companies rely on an internal team for software testing

Nearly all companies (98%) rely on internal testers for their software quality assurance. Specifically, 80% of responses rely on an internal team with external testing tools, while 89% leverage an internal team with internally developed tools. A smaller segment (26%) also incorporates external teams.



Which of the following does your company need to improve its ability to properly test software?



Despite widespread internal testing, companies identify several areas for improvement. The most pressing needs are more appropriately skilled testers (60%), better practices and success metrics (57%), and better testing tools (54%).

Interestingly, 51% also expressed a need for more direction from leadership and more time to devote to testing. A combined 82% indicated a need for either better testing tools or more skilled testers.

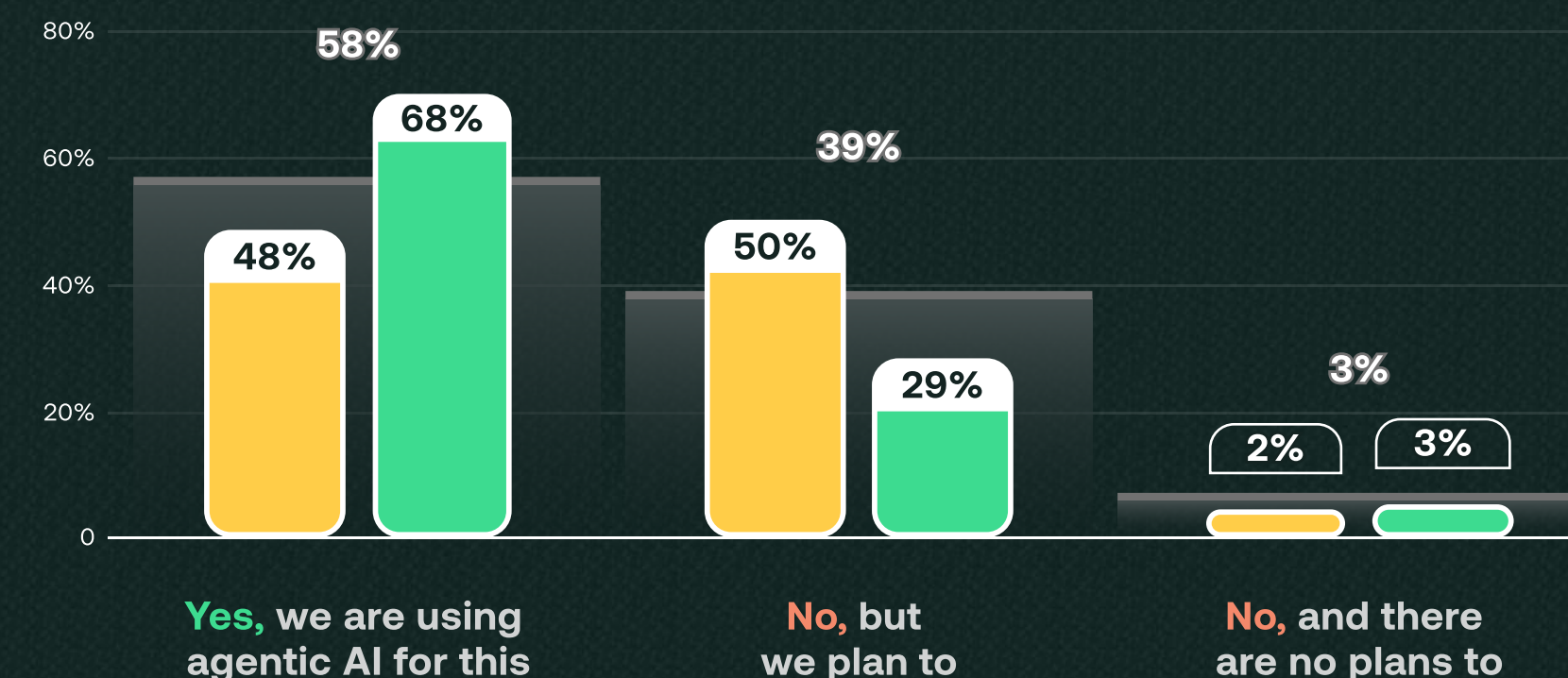
The data paints a clear picture of an industry that has embraced internal testing as the foundation of quality assurance, yet recognizes significant room for improvement. With over four-fifths of companies saying they need either enhanced testing tools or more skilled personnel, the future of software testing requires strategic investment in both human capital and technological resources.

Quality assurance in software testing extends beyond tools and talent. It demands organizational commitment alongside structured processes. Companies that address these multifaceted needs—combining skilled testers, advanced tools, clear metrics, and leadership support—will be best positioned to deliver superior software quality in an increasingly demanding market.

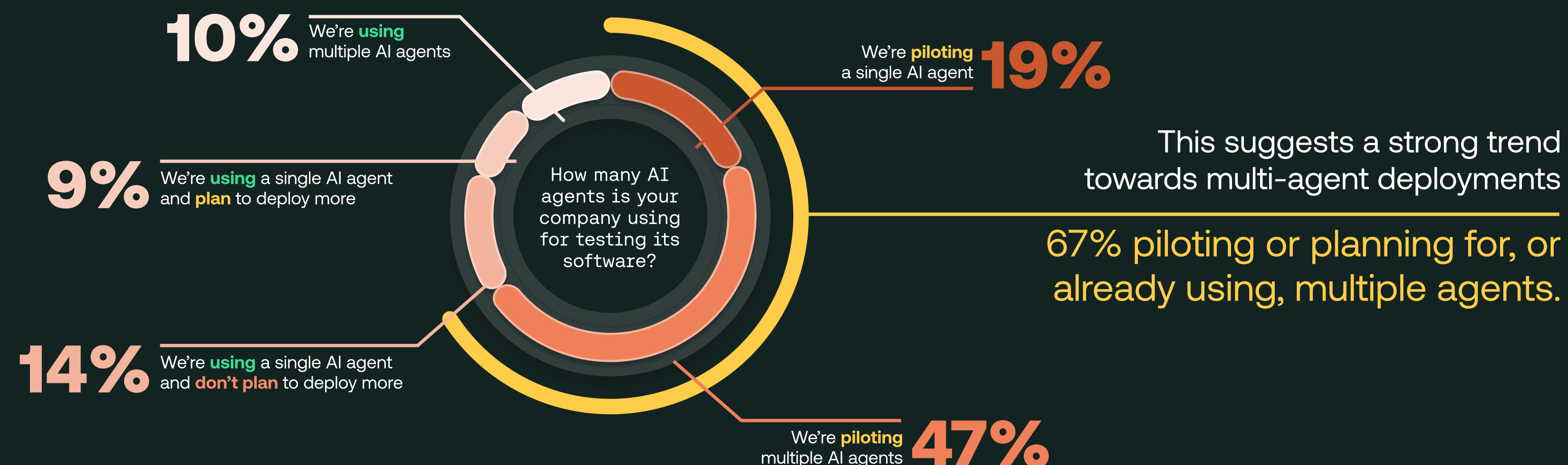
Agentic AI adoption trends

Among companies with internal testers, most (58%) already use agentic AI to support software testing. This adoption is notably higher among **engineering leaders (68%)** compared to **executives (48%)**, which suggests that executives may be misunderstanding agentic AI use on their teams. An additional 39% plan to use agentic AI, indicating that 97% of companies are either currently using or planning to use agentic AI for testing.

Thinking specifically of your company's work on testing its software
Is your company using agentic AI to support this work?



For those actively using agentic AI, a significant portion (47%) are piloting multiple AI agents, while 19% are piloting a single AI agent. Overall, 66% are in the piloting phase. When considering full deployment, **10% already use multiple AI agents**, and **9% are using a single AI agent and plan to deploy more**.

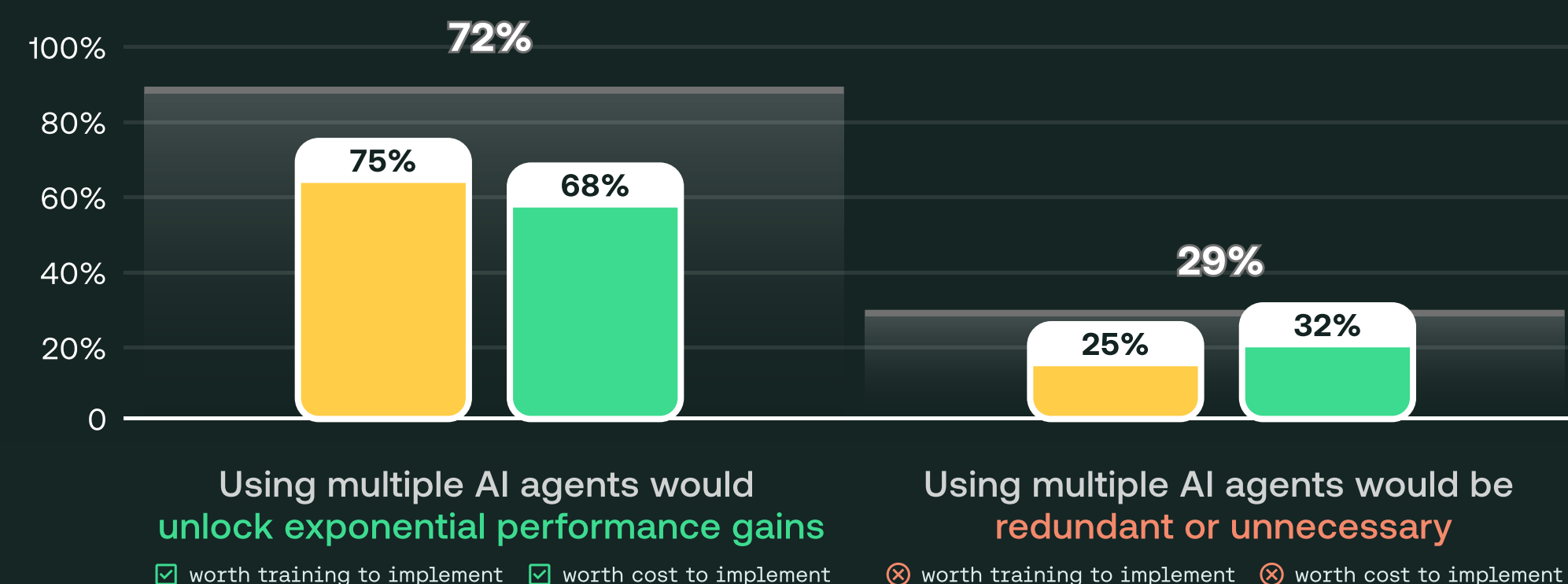


The sentiment towards multiple AI agents is overwhelmingly positive:

72% believe that using multiple AI agents would unlock exponential performance gains, making it worth the training and costs. This view is more prevalent among **executives (75%)** than **engineering leaders (68%)**.

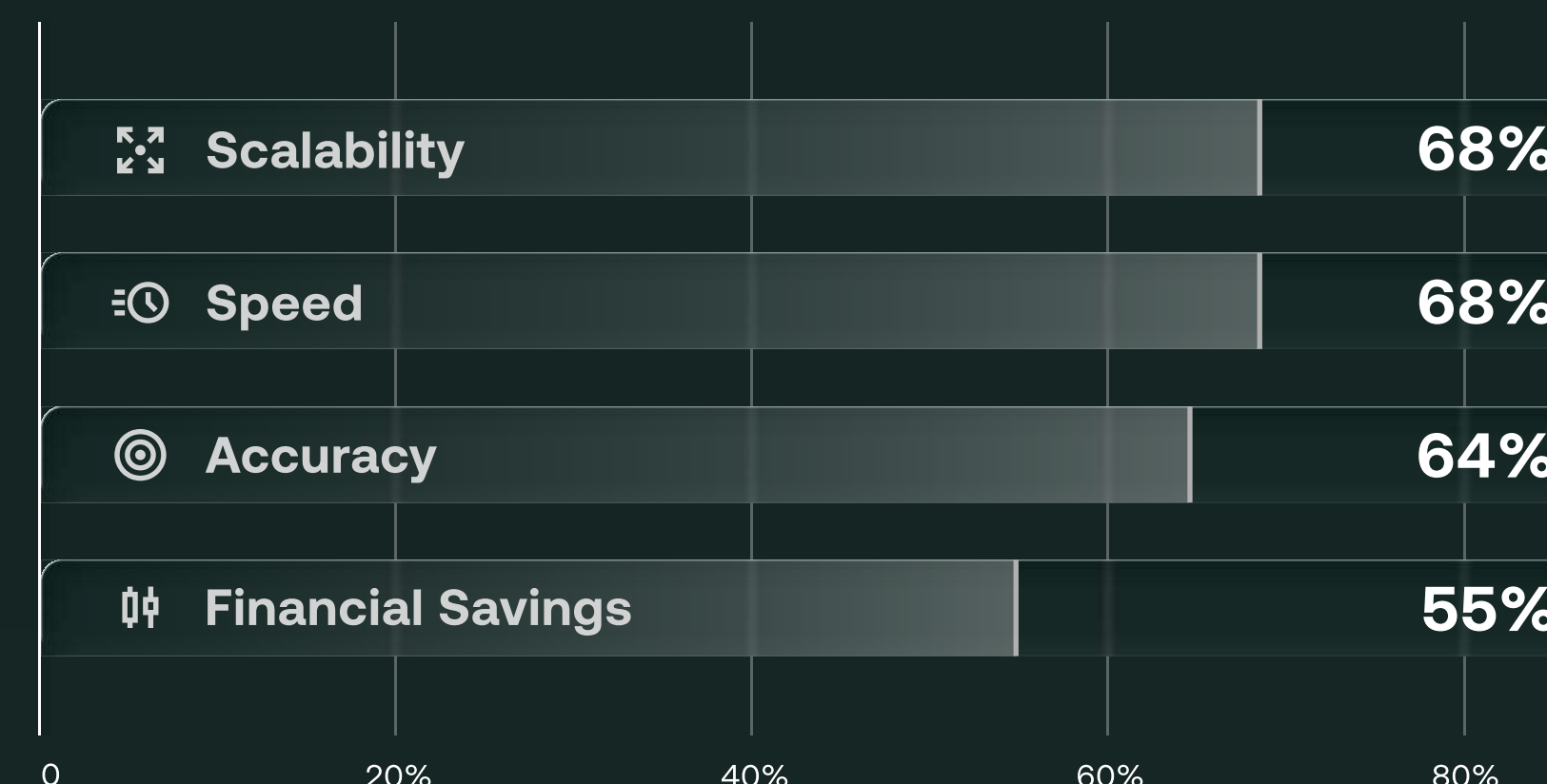
On the other hand, 29% responded that multiple agents would be redundant and not worth the extra cost and training to deploy.

Which better describes your opinion on using multiple AI agents (compared to one agent) for work processes?



The major perceived benefits of using agentic AI for software testing are scalability (68%), speed (68%), and accuracy (64%). Financial savings are also a significant benefit, cited by 55%.

Which of the following would you say are major benefits of using agentic AI for testing company software?



While the perceived benefits of agentic AI in software testing paint an optimistic picture, the current preference patterns reveal a disassociation between aspiration and practical reality. **Most current AI agents have consistently fallen short of their promised capabilities, making a majority-AI approach premature and potentially risky.**

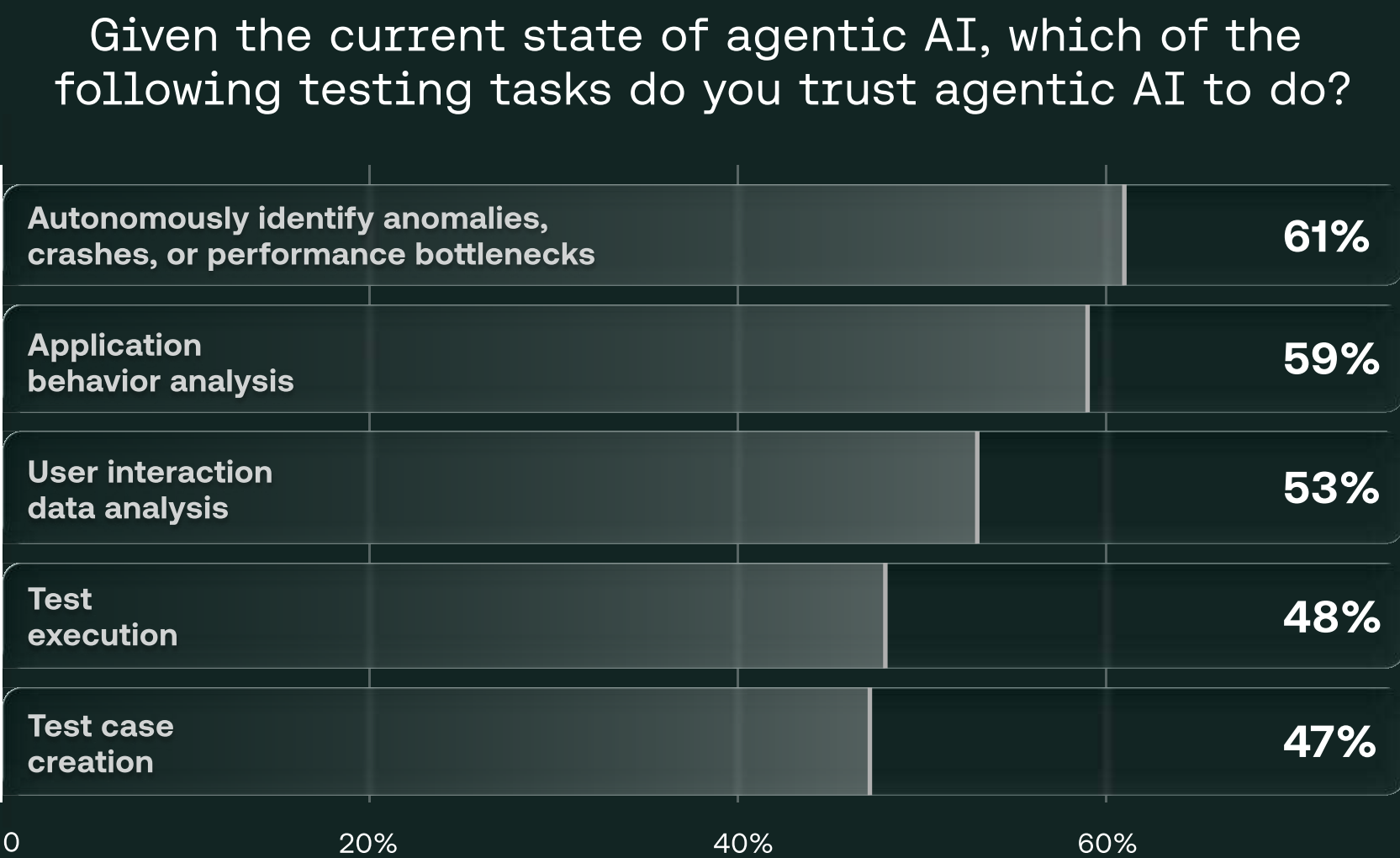
Trust is high and most say they have an AI strategy

A day probably doesn’t go by when CEOs and leaders are not thinking strategically about how their teams can best use AI as a business advantage.

Among active agentic users, 97% say they have a strategy in place. Specifically, 39% have a clear strategy and benchmarks, while 58% have a general strategy and are developing benchmarks. Only 3% have no strategy but plan to develop one.

Trust in agentic AI appears to be considerable:

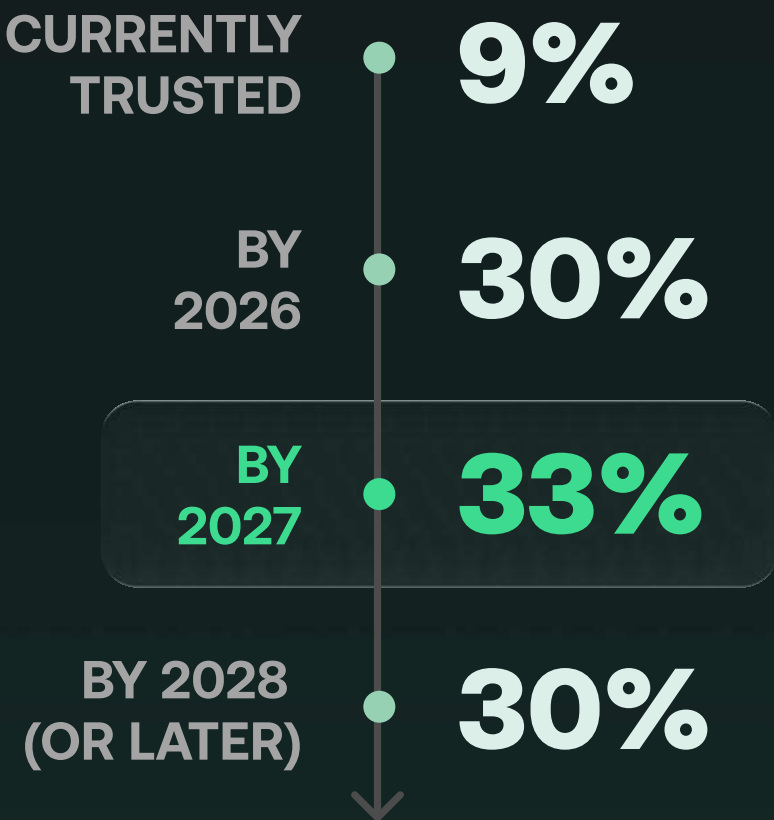
61% of respondents said that they trust agentic AI to autonomously identify anomalies, crashes, or performance bottlenecks, while 59% trust agentic AI for application behavior analysis during run time, 53% for user interaction data analysis, 48% for test execution, and 47% for test case creation. This may be a bit of an overly optimistic view because the AI abilities of current software testing tools fall short of reliably carrying out most of the tactics listed above.



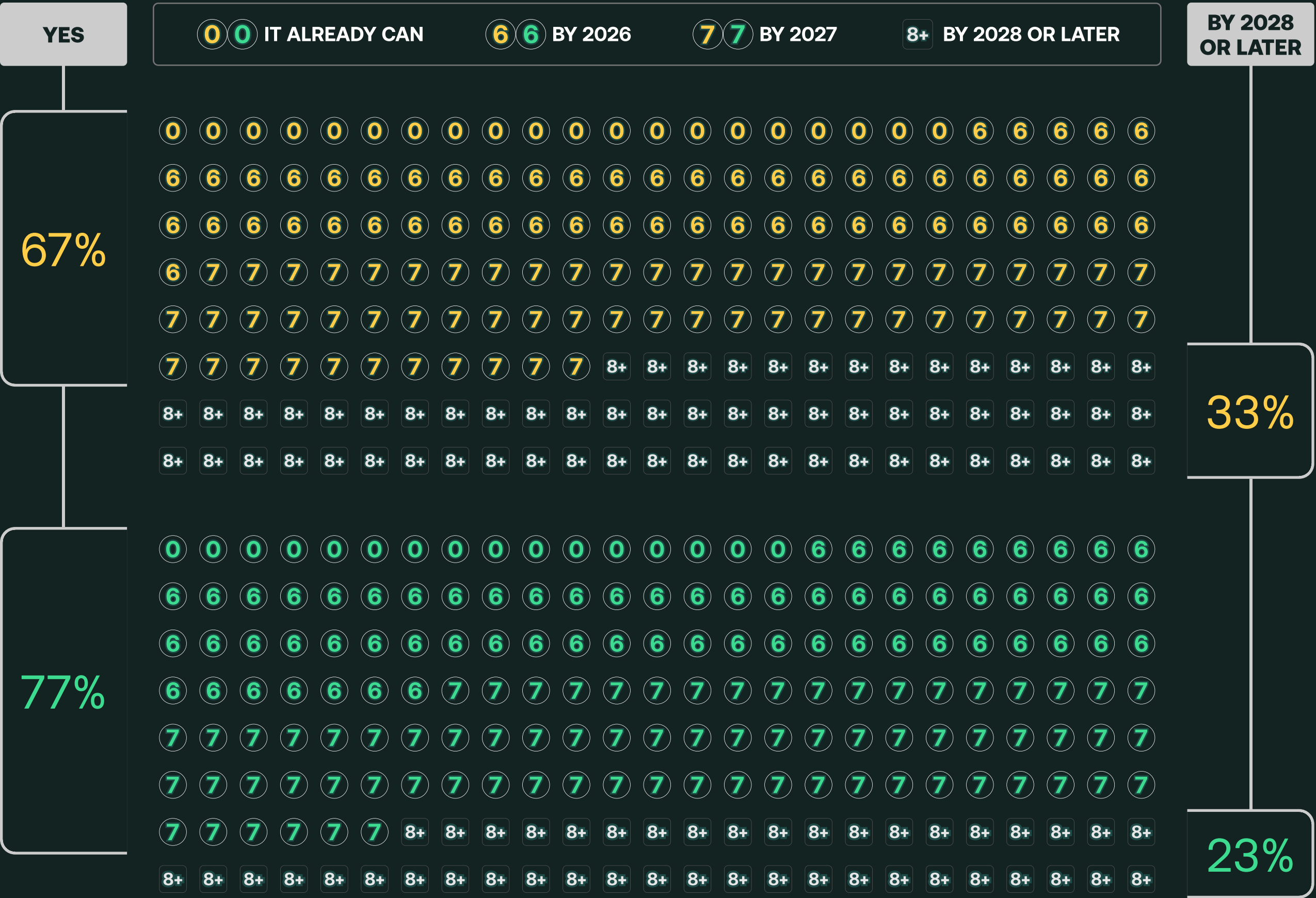
A combined 72% believe that agentic AI will be able to test software completely autonomously by 2027.

This includes 9% who think it's already possible, 30% who foresee it by 2026, and 33% by 2027. Engineering leaders (77%) are more optimistic about this timeline than executives (67%). This bullish sentiment reflects the persistent hype swirling about the media and social media platforms. It doesn't appear to be going away anytime soon.

How soon do you think agentic AI can be trusted to test software completely autonomously with no human intervention or oversight?



Do you think agentic AI can be trusted to test software completely autonomously with no human intervention or oversight by 2027



However, concerns about data and security access persist.

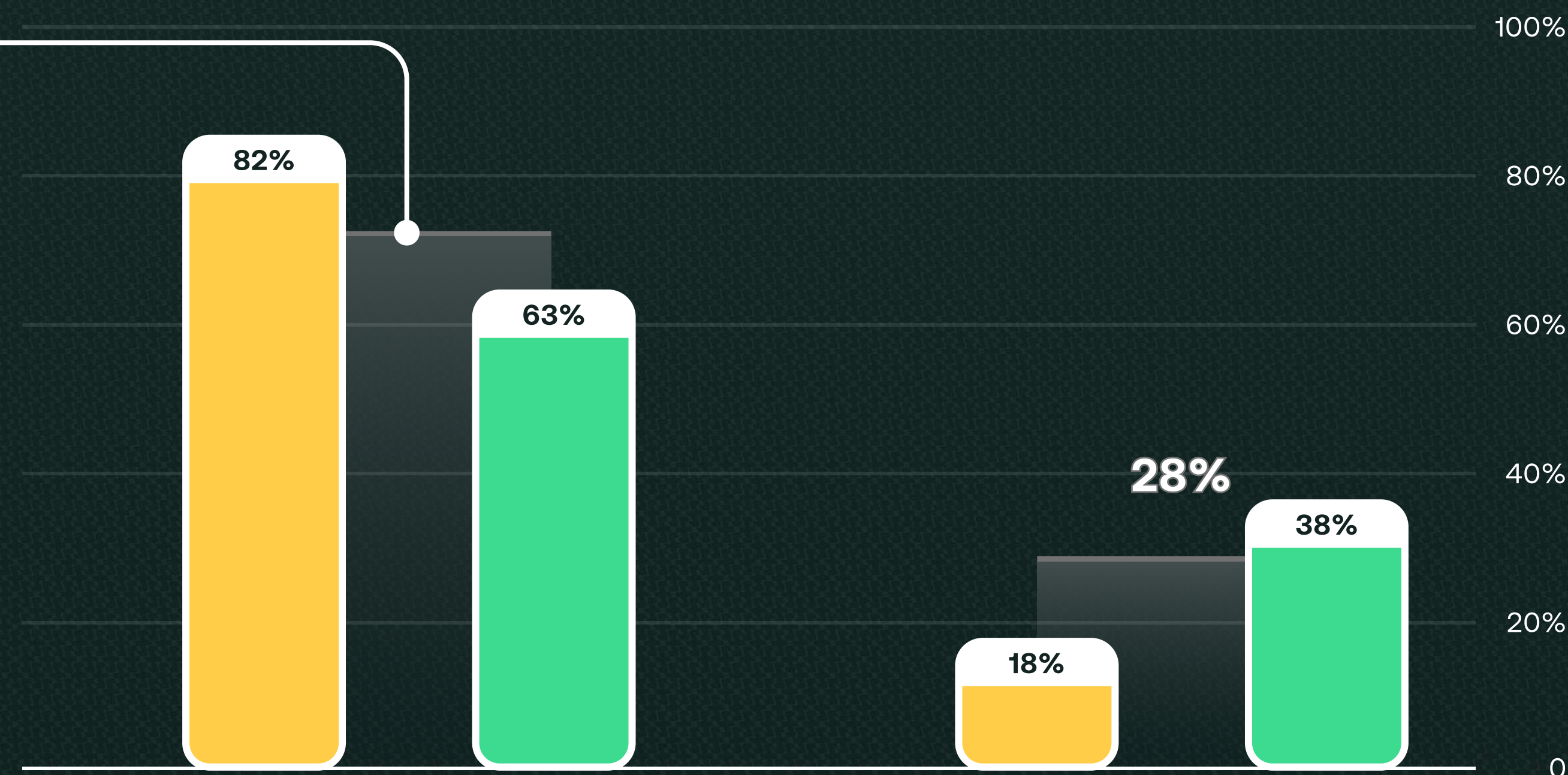
A majority (72%) are more nervous about granting AI agents full company data and security access than falling behind competitors who do.

This concern is significantly higher among **executives (82%)** than **engineering leaders (63%)**.

Which makes you more nervous?

Granting AI agents full company data and security access

Falling behind competitors who grant AI agents full company data and security access

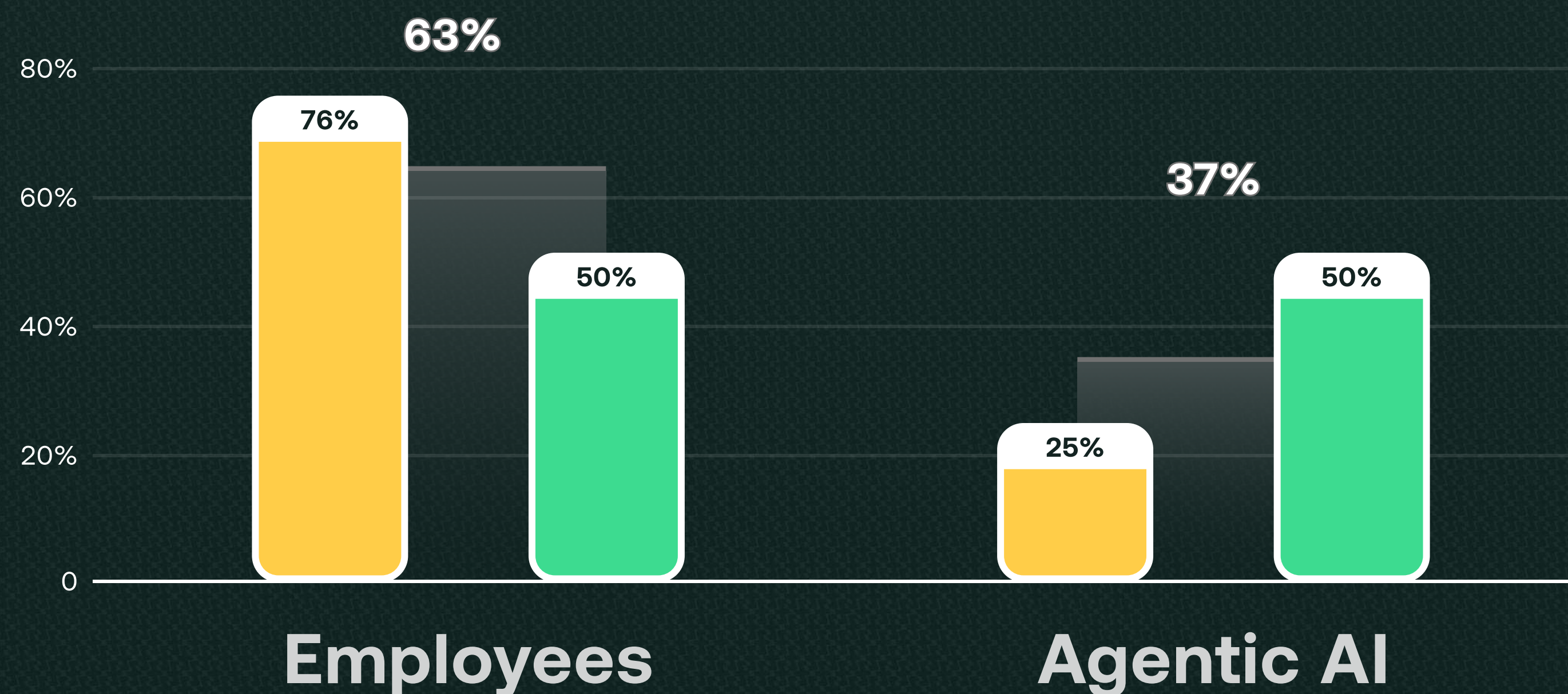


Regulation, Compliance, and AI

When it comes to meeting regulatory requirements for compliance, 63% trust employees more than agentic AI (37%).

This trust in employees is particularly strong among executives (76%), while engineering leaders are evenly split

Who do you trust more to accurately meet regulatory requirements in your industry for compliance?



The real question may be:

Even if AI could fully autonomously write software tests, *should it?*

Despite these concerns, 88% believe their company's top leadership fully understands the risks of introducing agentic AI to software testing. **Engineering leaders (91%)** feel slightly more confident in leadership's understanding than **Executives (86%)**.

While organizations demonstrate remarkable confidence in their leadership's risk assessment capabilities and maintain ambitious timelines for fully autonomous testing, the tension between competitive pressure and data security remains a defining challenge.

The notable gap between engineering leaders and executives—particularly regarding security concerns and timeline expectations—suggests that successful agentic AI implementation will require continued alignment across organizational levels. Ultimately, companies that can balance their strategic ambitions with robust security frameworks while fostering cross-functional consensus will be best positioned to reap the potential benefits of agentic AI.

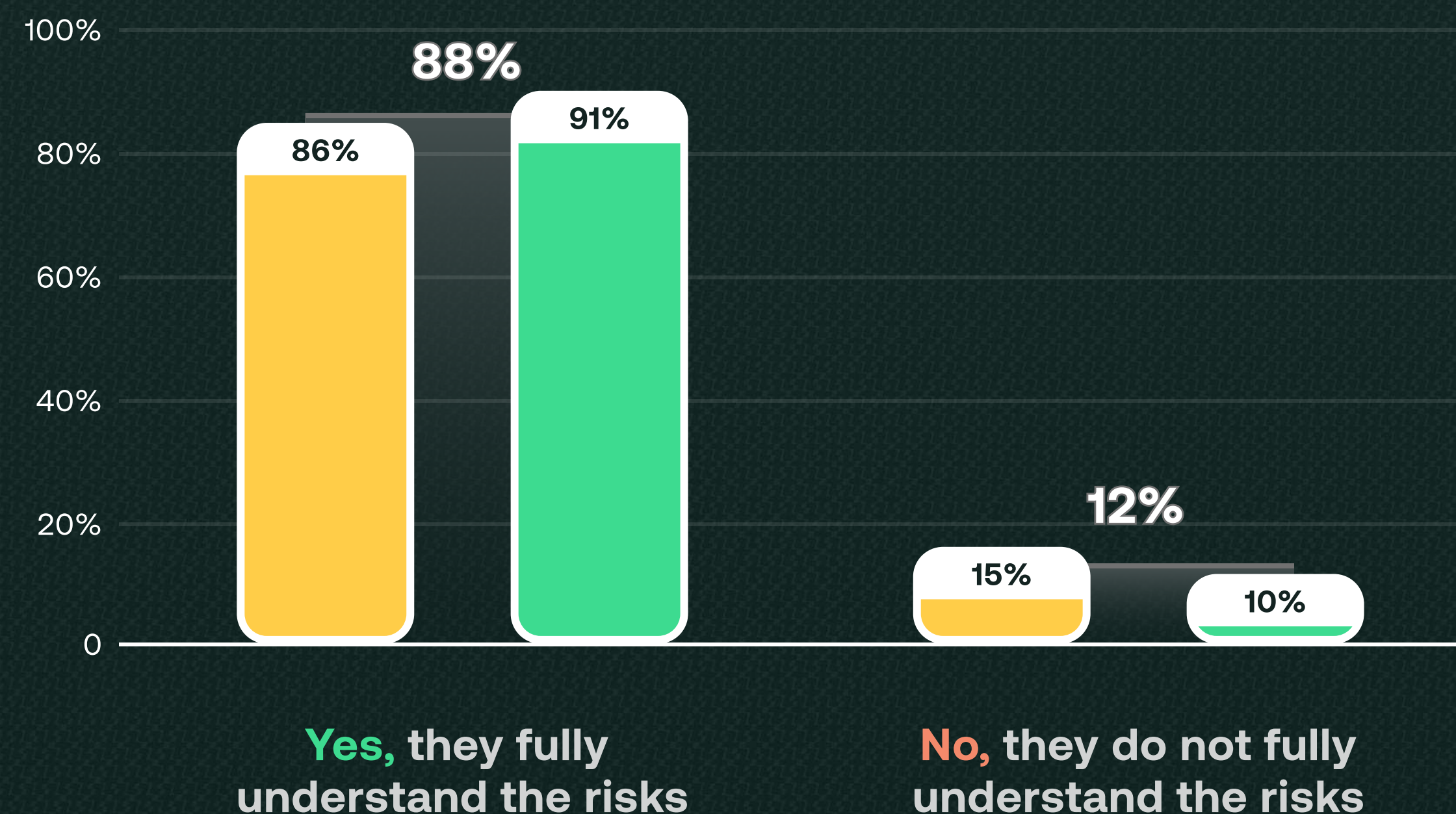
LEGEND

■ Executives

■ Engineering Leaders

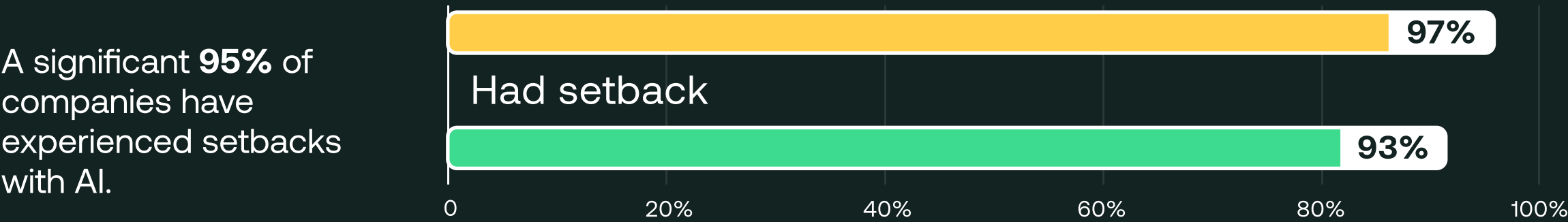
■ Totals

Does your company's top leadership fully understand the risks of introducing agentic AI to software testing?

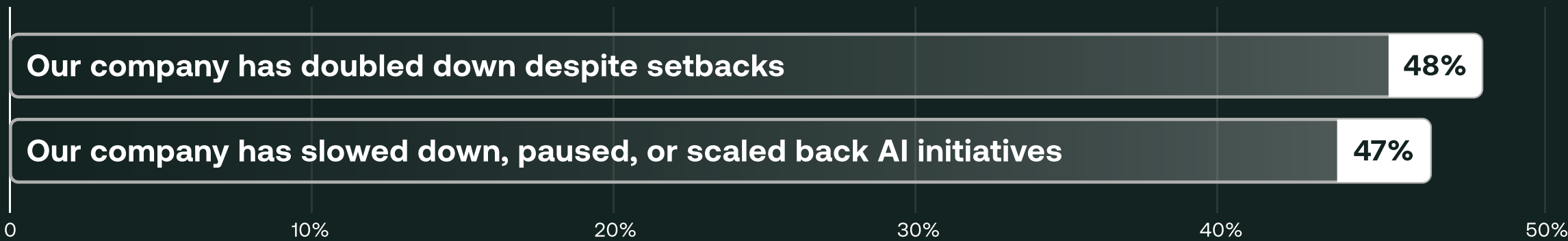


Response to AI setbacks

The promise of agentic AI's transformative benefits has been met with an unexpected reality check. The path to success is far from straightforward. Understanding how organizations respond to these obstacles provides crucial insights into the resilience required for long-term success.



When faced with setbacks or negative outcomes from generative AI tools, the response is pretty well split between respondents. Half of the companies (48%) have doubled down on their initiatives, with the other half (47%) slowing down or even pausing AI initiatives.



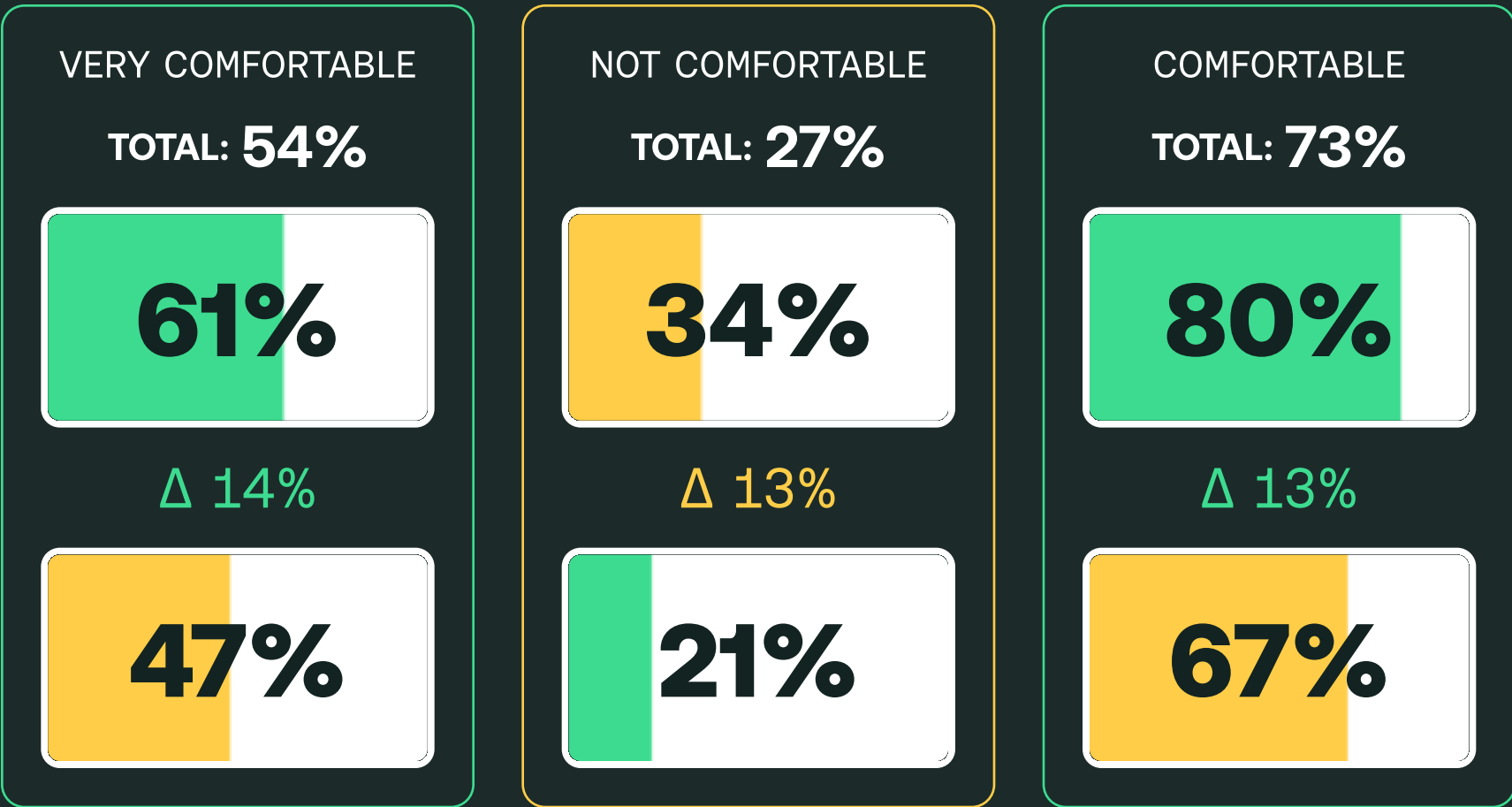
Organizations that view setbacks as learning opportunities and maintain their commitment may ultimately gain competitive advantages, while those that pause risk falling behind others.

Coming soon to a department near you:

AI agents determining layoffs?

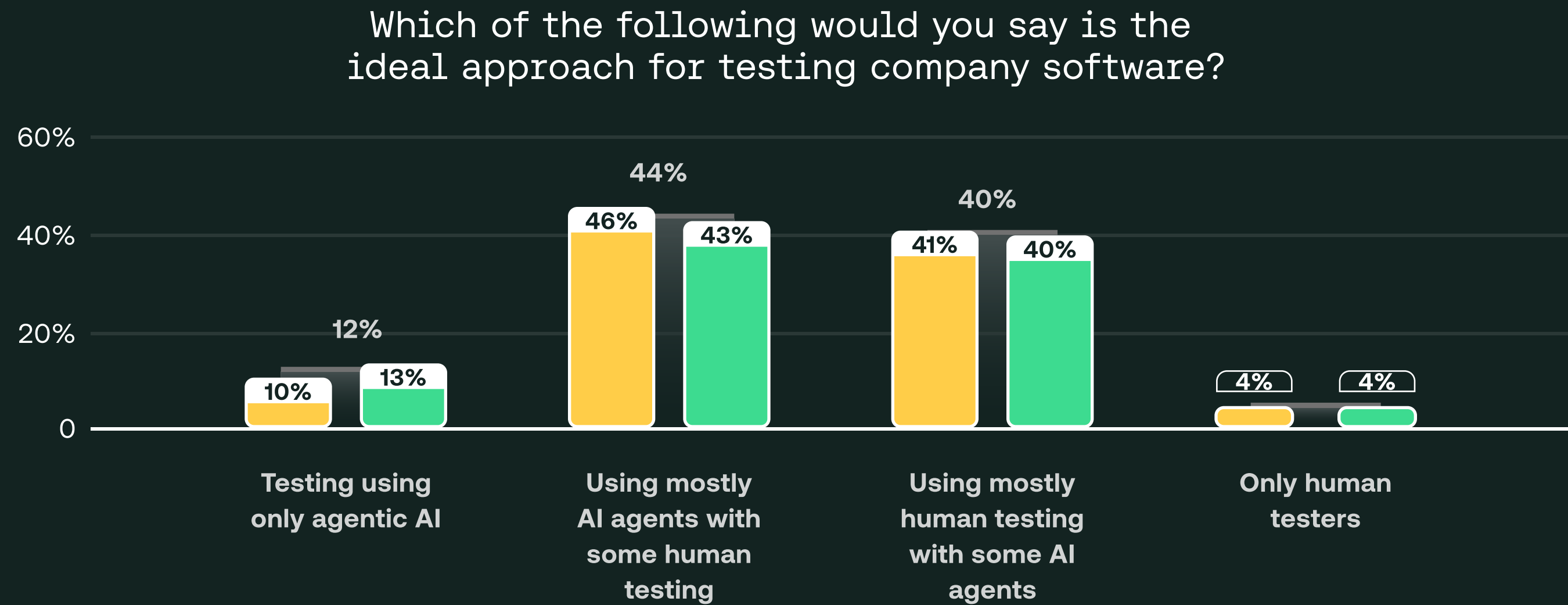
54% of survey respondents said they would be very or extremely comfortable with AI agents determining job cuts. While 27% are not comfortable, engineering leaders seem to be more comfortable than executives with the idea: 61% vs. 47%, respectively.

How comfortable would you feel with AI agents determining job cuts, including eliminating top leadership positions?

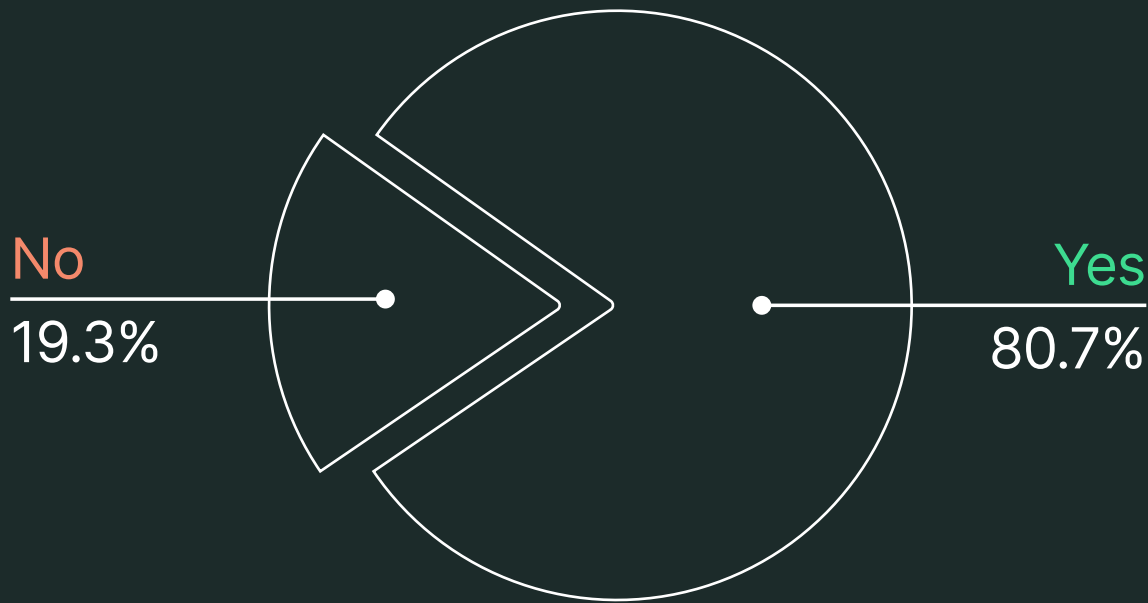


Striking the winning balance

When asked about the ideal approach for software testing, a mixed model prevails. While only 12% advocate for testing using only agentic AI, a substantial **44% prefer mostly AI agents with some human testing**. Another **40% favor mostly human testing with some AI agents**. This indicates a strong preference for a hybrid approach, with 85% opting for a mix of human and AI testing. Overall, 56% believe the ideal approach involves only or mostly AI agents.



Should companies publicly disclose when software has **only been autonomously tested**?



The overwhelming leaning towards hybrid models demonstrates that organizations recognize the irreplaceable value of human judgment and oversight in critical testing scenarios.

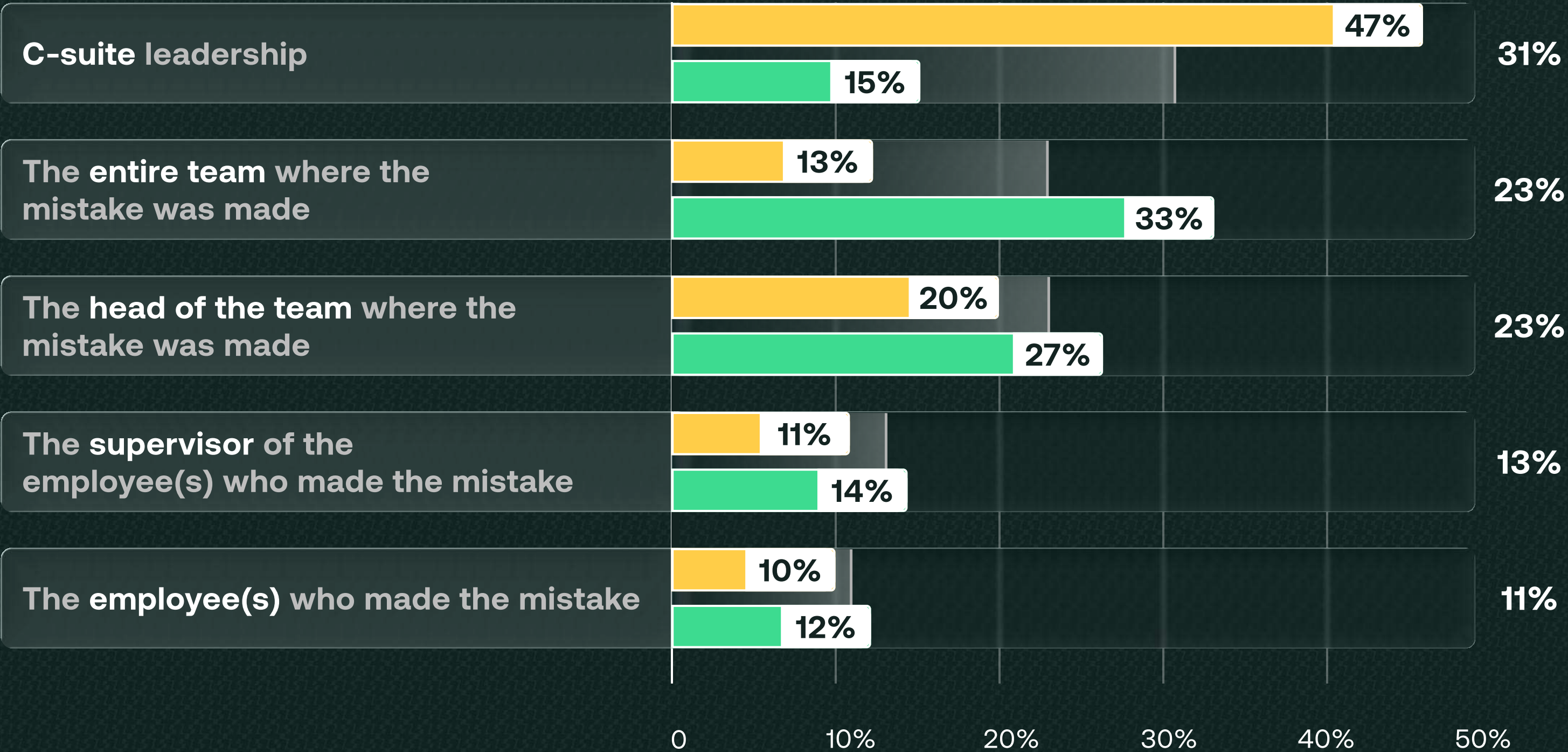
Until AI agents can prove consistent reliability and accuracy that match their theoretical benefits, the evidence strongly supports the hybrid approach that most organizations are wisely adopting—leveraging AI's strengths while maintaining human expertise as the essential foundation of quality assurance.

When in doubt, blame a robot?

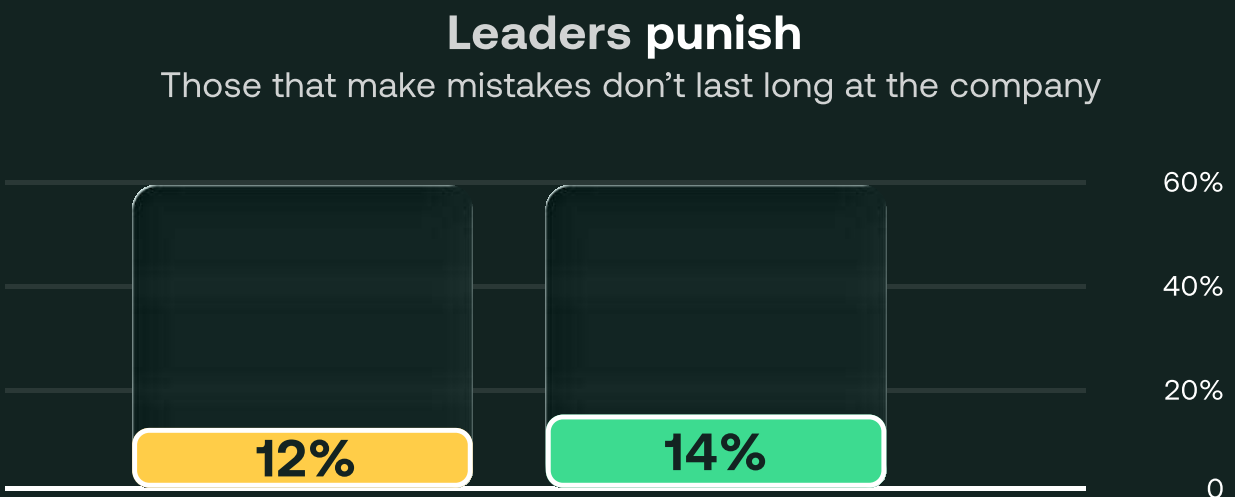
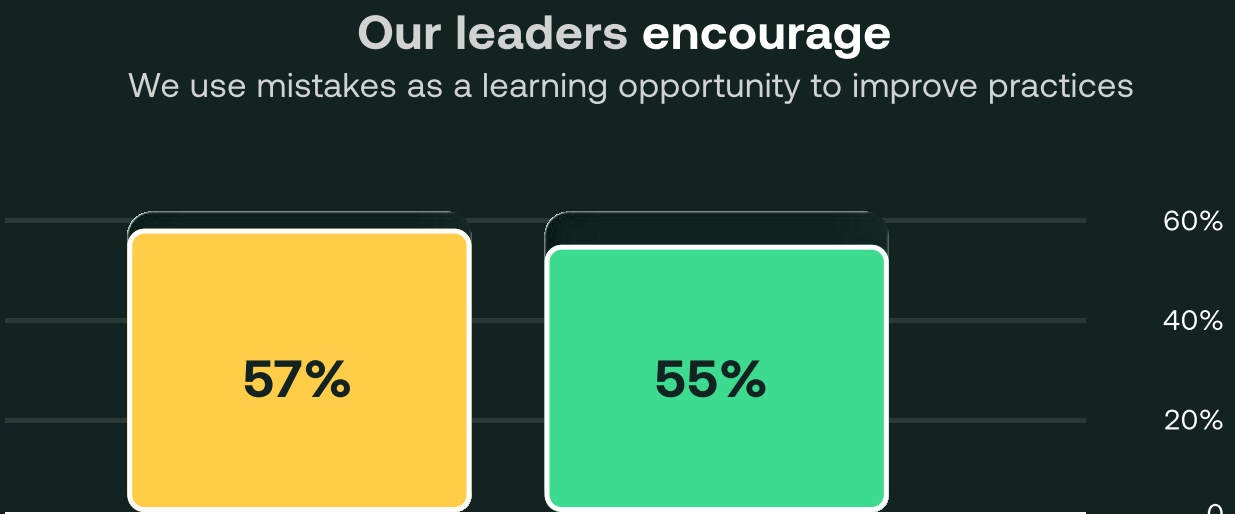
Who is really to blame when AI messes up?

If a significant mistake occurs, C-suite leadership is seen as bearing the primary responsibility by 31% of respondents, with executives (47%) being significantly more likely to assign blame to C-suite than engineering leaders (15%). Other significant bearers of responsibility include the entire team (23%) and the head of the team (23%).

If someone were to make a major mistake at your company, who would ultimately bear the primary responsibility?

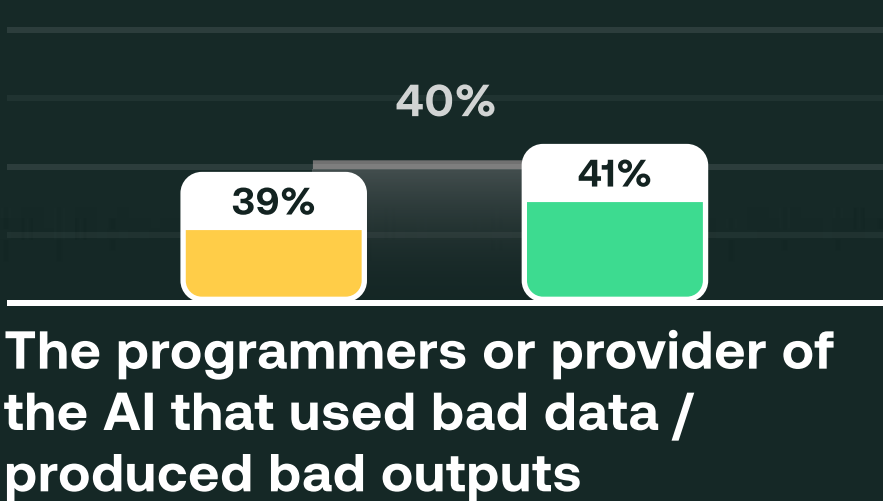
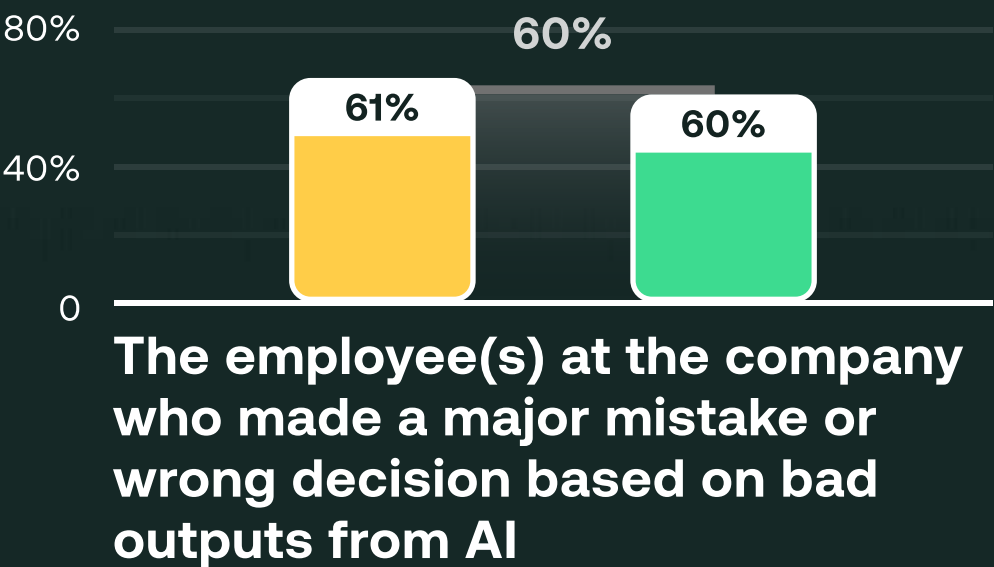


56% of leaders encourage employees to use mistakes as a learning opportunity. However, 32% report inconsistent approaches, and 13% indicate punishment.



If a major mistake or wrong decision is made using bad AI outputs, the brunt of the blame from top leadership is seen as falling on the employee(s) at the company who made a major mistake or wrong decision based on bad AI outputs (60%) rather than the programmers or provider of the AI (40%).

If someone were to make a **major mistake** or **wrong decision** at your company using bad outputs from AI, who would get the brunt of the blame from top leadership?



The stark contrast between executives' reliance on human judgment for compliance versus engineering leaders' balanced perspective highlights a foundational misalignment in AI readiness across some organizational hierarchies.

The tendency to hold employees accountable for AI-generated errors (60%) while leadership remains disconnected mainly from testing realities creates a precarious environment where those closest to the technology bear disproportionate responsibility for its failures.

This combination of leadership knowledge gaps, inconsistent accountability practices, and the expectation that employees should navigate AI limitations without adequate support from above suggests that many organizations are ill-equipped to integrate agentic AI into their critical testing processes successfully.

How to move forward in an agentic world

What does all of this survey data really mean to you? We've compiled some tips and actionable advice for you as next steps in adopting agentic AI at your organization:

GENERAL TESTING STRATEGIES

- 1 First and foremost, start addressing the leadership knowledge gap.**
Most respondents indicate that that top leaders don't understand software testing. That is a problem. Invest in technical education for executives before making major AI investments. Create cross-functional alignment between engineering leaders and executives, particularly around security concerns and timeline expectations, as their perspectives differ significantly.
- 2 Pilot a few agents, but maintain realistic timelines.**
While 72% believe full autonomous testing is achievable by 2027, focus on specific use cases where trust is highest: anomaly detection (61%) and runtime behavior analysis (59%).
- 3 Invest in both skilled testers and improved tools, as 82% of companies need one or both.**
Your AI strategy should complement, not replace, human expertise—at least until the technology proves it can consistently deliver on its promises.
- 4 Develop clear accountability frameworks before deployment.**
Since 60% of blame for AI errors falls on employees rather than the technology, establish transparent protocols for AI-generated mistakes that protect your team while encouraging innovation. Consider that 72% of organizations prioritize data security over competitive advantage—ensure your security framework matches this cautious approach.
- 5 Falling behind competitors in AI is a very real threat for companies.**
But that doesn't mean you should let hype overtake common sense and reality. There are lots of conversations swirling around, but only you know what's best for your team and company in a given point of time.

The survey results expose a fundamental tension between ambition and apprehension.

While many believe multiple AI agents will deliver exponential performance gains and expect fully autonomous testing by 2027, an equal percentage remain nervous about granting AI agents comprehensive data and security access. This dichotomy reflects the complex reality of AI adoption: organizations recognize the transformative potential while grappling with legitimate security and control concerns. Most know where they want to go but nobody has the map.

Explore how Sauce Labs brings AI into its continuous quality platform on the [Sauce AI hub](#). >>



VIRTUAL
DEVICE CLOUD



REAL
DEVICE CLOUD



SAUCE
VISUAL



MOBILE APP
DISTRIBUTION



CRASH & ERROR
REPORTING



ADVISORY
SERVICES



Sauce Labs is the leading cloud-hosted platform for automated testing of web and mobile applications, enabling fast delivery of high-quality software across the development lifecycle. Founded by the creators of Selenium, Sauce Labs has been the testing leader for over 15 years and now runs over 1 billion tests annually. Trusted by Fortune 500 companies like Toyota, Walmart, Verizon, and Fidelity Investments, its scalable, secure platform supports testing across thousands of operating systems, browsers, and devices while meeting the highest compliance standards.