FORRESTER®

Connecting AI Output Quality To Business Outcomes

AI Quality Is Multifaceted And Requires A Holistic Approach

A FORRESTER CONSULTING THOUGHT LEADERSHIP PAPER COMMISSIONED BY MICROSOFT, JULY 2026



Table Of Contents

3 Executive Summary

4 Key Findings

5 Expectations For AI Are Increasing

8 Accuracy, Speed, And Trust Drive AI Output Quality

14 Improve AI Output Quality With A Holistic Approach

16 Key Recommendations

17 Appendix



Project Team:

Chris Taylor, Principal Consultant

Contributing Research:

Forrester's Future Of Work research group

ABOUT FORRESTER CONSULTING

Forrester provides independent and objective research-based consulting to help leaders deliver key outcomes. Fueled by our customer-obsessed research, Forrester's seasoned consultants partner with leaders to execute their specific priorities using a unique engagement model that ensures lasting impact. For more information, visit forrester.com/consulting.

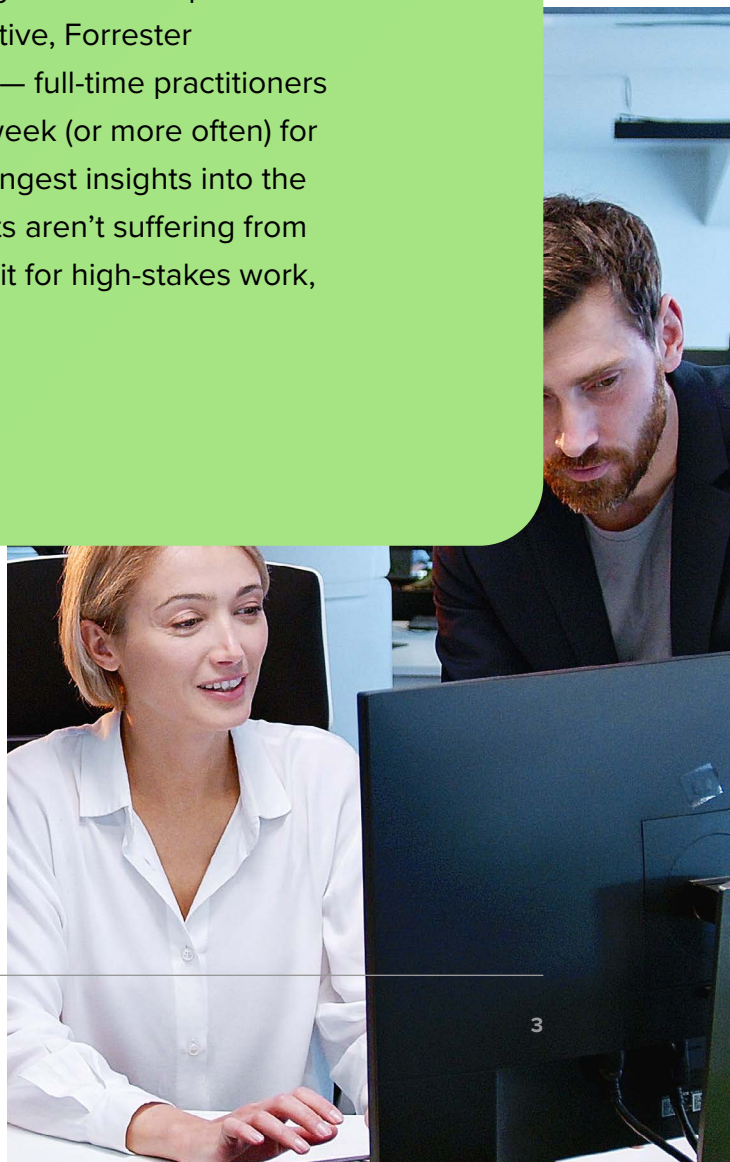
© Forrester Research, Inc. All rights reserved. Unauthorized reproduction is strictly prohibited. Information is based on best available resources. Opinions reflect judgment at the time and are subject to change. Forrester®, Technographics®, Forrester Wave, and Total Economic Impact are trademarks of Forrester Research, Inc. All other trademarks are the property of their respective companies. [E-70447]

Executive Summary

AI advancements are happening at a record pace, and employee, consumer, and business adoption continues to surge. However, despite growing use, few enterprises have yet to turn early AI innovation into significant business impact.[1] Underlying the entire AI opportunity, and its challenges, is the premise of AI output quality. If users lack confidence in the quality of outputs, AI will undoubtedly fail to live up to expectations.

There are multiple tenets of quality to consider when evaluating AI output, particularly accuracy, speed, and trust. If organizational leaders want to drive improvements in AI output quality, they must weigh each. As business leaders build stronger foundations for enabling high-quality AI output, they will be in a better position to connect AI use with key business metrics and improvements.

In June 2026, Microsoft commissioned Forrester Consulting to evaluate how organizations and their employees are thinking about AI output quality and how to improve it. To achieve this objective, Forrester conducted an online survey of 608 AI power users — full-time practitioners and business leaders who use AI multiple times a week (or more often) for work. The focus on AI power users provide the strongest insights into the actual state of AI quality because these respondents aren't suffering from lack of skill: They engage with AI regularly and use it for high-stakes work, for which quality output is critical.



Key Findings

Even AI power users have AI output quality concerns. The majority of surveyed AI power users use AI for more than 50% of business tasks, yet 49% have concerns about AI output quality, and 40% say those concerns lead them to use AI less often than they or their employers would like.

AI output quality includes aspects of accuracy, speed, and trust. Generating a high-quality AI output is about more than just not having fake data. Accuracy is important but must be considered along with the time it takes to achieve desired outcomes. Additionally, trust in AI output (e.g., data coming from the right sources, having the right permissions) is a critical prerequisite for many users to even use AI in the first place.

Improving AI output quality requires enterprisewide training. More than 40% of respondents reported that their organization is investing in more organizationwide training and learning initiatives to improve AI competency. Output quality is likely to improve as employees' AI operating skills increase, such as learning how to better prompt or redirect a model, knowing how and when to verify output, and learning key signals of untrustworthy results.

Expectations For AI Are Increasing

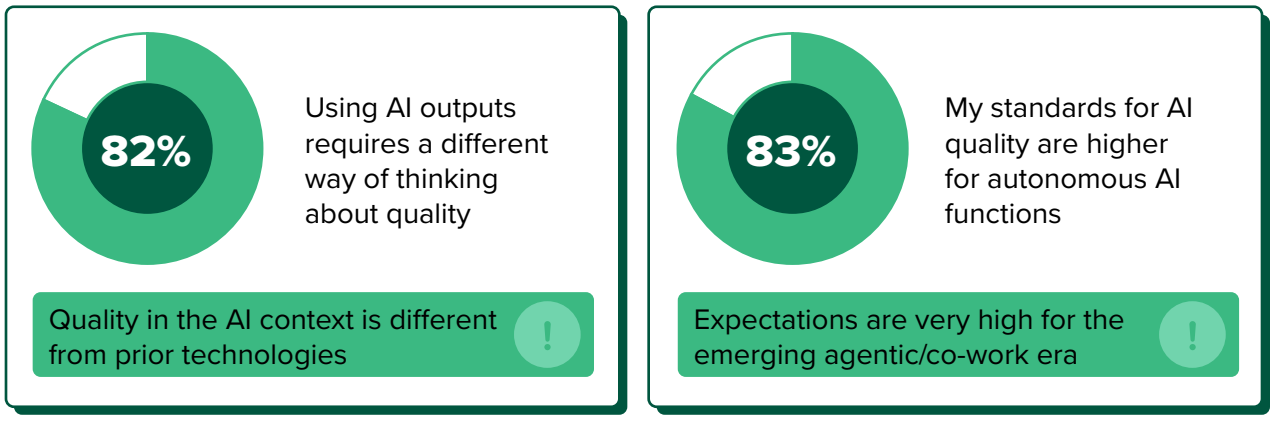
As with every investment, business leaders want to be able to link AI investments to positive business outcomes. Doing so will allow them to better justify ongoing investments and focus on the right capabilities and improvements. Our survey found that most leaders are hoping to track AI's impact on employee experience (57%), customer experience (56%), individual productivity (56%), and risk reduction (51%).

However, before organizations can begin to measure the impact of AI, they first need to be confident that AI is creating the desired output. Since many AI use cases today include generative AI functions, standards for AI are different than other employee-created outputs: 82% of respondents agreed that AI outputs require a different way of thinking about quality, and 83% agreed that their standards for AI quality are even higher for autonomous AI functions (see Figure 1). Therefore, the first step toward being able to demonstrate the business value of AI is to define expectations for AI output quality to ensure outputs are valuable.

FIGURE 1

AI Output Quality Is Top Of Mind

(Showing combined "Strongly agree" and "Somewhat agree" responses)



Base: 608 global AI power users across various industries from a mix of practitioner and decision-maker roles
Source: Forrester's Q2 2026 AI Output Quality Survey [E-70447]

AI POWER USERS OFFER THE CLEAREST LOOK INTO CURRENT AI OUTPUT REALITIES

The people best suited to discuss AI output quality are those who use it frequently, understand how it works, know its limitations, and are trained on how to prompt properly to achieve desired outcomes. Forrester defines this set of skills as AIQ (the Artificial Intelligence Quotient), a framework that measures the readiness of individuals, teams, and organizations to adapt to, collaborate with, trust, and generate business results from generative AI.[2]

Ninety-six percent of respondents from our survey score Medium AIQ to High AIQ and have high rates of AI use: 65% consider AI essential to their day-to-day job functions, 57% use it multiple times a day, and 57% use it for more than half of their work tasks. The heavy AI use by this group of respondents makes them ideal to speak to the importance and challenges of AI output quality: They aren't lacking in skills, they engage with AI regularly, and they understand its strengths and weaknesses. For this study, we'll refer to this group as AI power users.

HEAVY AI USE DOES NOT IMPLY FULL CONFIDENCE IN AI OUTPUT

One of the more surprising findings from the survey of AI power users is that 49% said they have concerns about the quality of AI output. Their concerns are significant enough that 40% stated they use AI less often than they'd like, and 41% said they do so less often than their employer would like them to, because of quality concerns (see Figure 2). This aligns with other findings showing that across various tasks, less than half (44%) of users on average are "Highly satisfied" with AI output results. Tasks include activities such as searching large amounts of data quickly, completing complex tasks, automating repetitive tasks, and uncovering deeper insights.

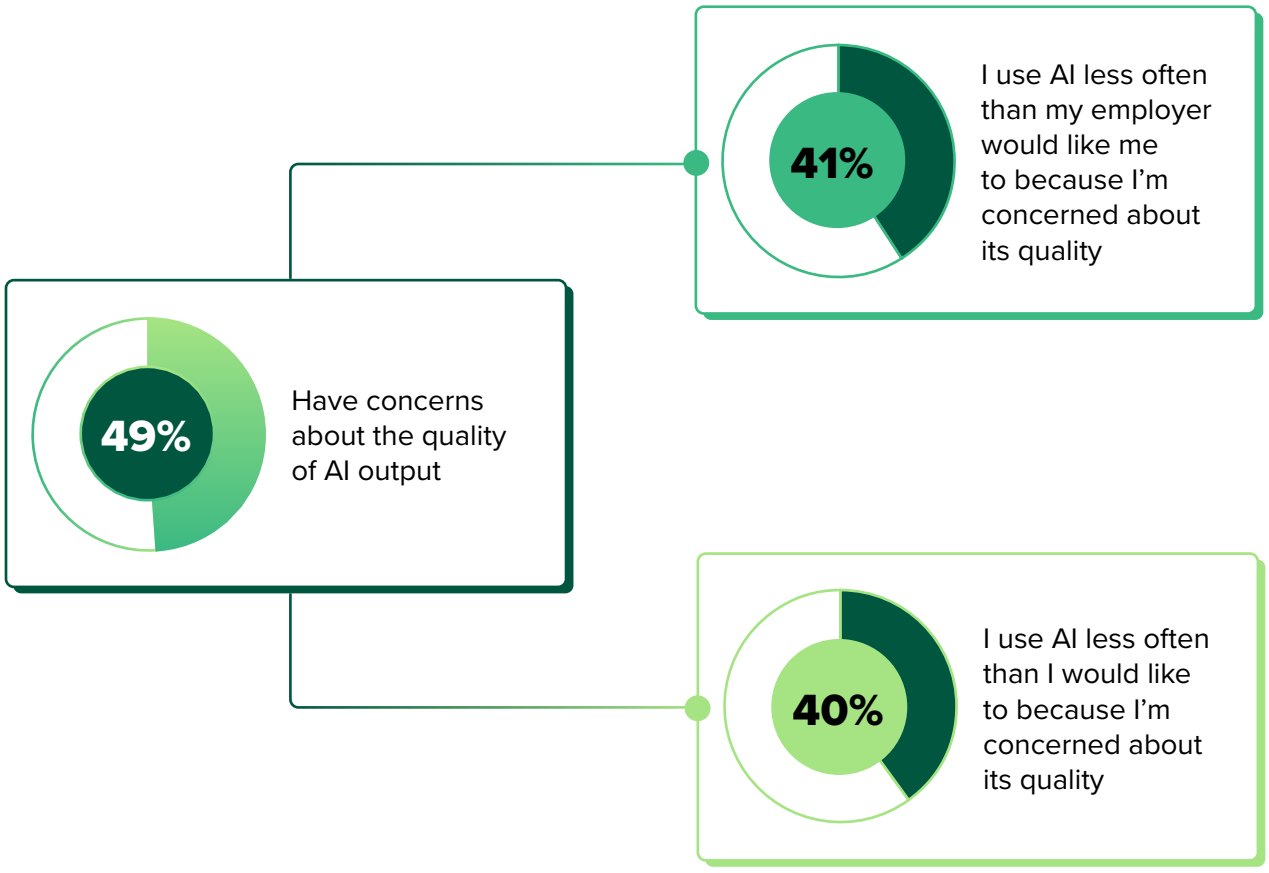
When quality expectations go unmet, it can lead to challenges with how to best complete tasks. Most users (65%) will attempt to reprompt with better instructions. This can be a common behavior as each new prompt may provide incremental improvements to the output. However, the need for frequent reprompting becomes problematic when users are forced to do so because the first answer hallucinated, missed context, or was overconfident. Another common reaction to poor-quality output for nearly

50% of respondents is to complete the task manually or switch to another AI tool. The survey didn't specify whether this abandonment takes place before or after reprompting attempts, but either way, switching to other AI tools may involve users moving to unsanctioned, public AI tools, which can create security risks. Also, by switching tools, AI will miss some of the business context surrounding the prompt, which is essential for quality output. Forty percent of AI power users stated that AI misunderstanding business context is a common reason AI output fails to deliver satisfactory results.

FIGURE 2

Half Of Surveyed AI Power Users Have Concerns About AI Output Quality

(Showing combined "Agree" and "Strongly agree" responses)



Base: 608 global AI power users across various industries from a mix of practitioner and decision-maker roles
Source: Forrester's Q2 2026 AI Output Quality Survey [E-70447]

Accuracy, Speed, And Trust Drive AI Output Quality

Current concerns about AI output quality warrant a deeper dive. AI output quality is not a binary good or bad. Instead, AI output quality should be viewed on a spectrum comprising multiple attributes, each with similar importance. When we asked AI power users to select which attributes they most commonly consider when evaluating AI output quality, responses were fairly evenly distributed across three primary attributes (see Figure 3):

- **Accuracy (72%).** AI output is error-free, grounded in reliable sources/ context, verifiable, and relevant to the initial prompt.
- **Speed (68%).** The time it takes a user to interact, iterate, and reach a satisfactory outcome, which also includes consistency and reliability for users to get the same results over time with similar prompts. For this definition, speed is not about latency or processing speed.
- **Trust (73%).** The belief that AI use is secure and complies with privacy policies and regulations, which also includes confidence that data being accessed by AI systems is from allowable sources with valid permissions.

AI power users harbor significant concerns about the potential negative business impacts stemming from poor-quality AI output, including the loss of client trust (37%), damage to company reputation (35%), and AI-derived errors leading to lost opportunities (34%). Fewer AI power users have actually experienced these negative impacts directly: 8% have lost customer trust, 7% have seen damage to company reputation, and 10% have lost opportunities (see Figure 4). Given AI power users' concerns, they're likely exploring preventive solutions to mitigate these negative impacts.

To understand the root causes of these concerns, it's important to take a deeper look at how power users are approaching these three primary attributes of AI output quality.

FIGURE 3

Accuracy, Speed, And Trust Drive AI Output Quality

Percentage of respondents selecting at least one option

ACCURACY

Verifiability of results (e.g., ability to track results back to clear sources)

Accuracy of output (e.g., the number of factual errors and/or hallucinations present)

72%



SPEED

Consistency and repeatability of output (e.g., we receive the same results over time with similar prompts)

Speed of the process (e.g., ability to get things right the first time or with few iterations)

68%



TRUST

Trust that the output is secure and compliant (e.g., not violating our best practices or relevant laws)

Trust that the output comes from allowable sources (e.g., AI system accesses data with valid permissions)

73%



CONTEXTUAL RELEVANCE

Contextual relevance of output (e.g., it makes sense with relation to current need)

39%



EXPLAINABILITY

Explainability (i.e., AI system can explain how it arrived at its output)

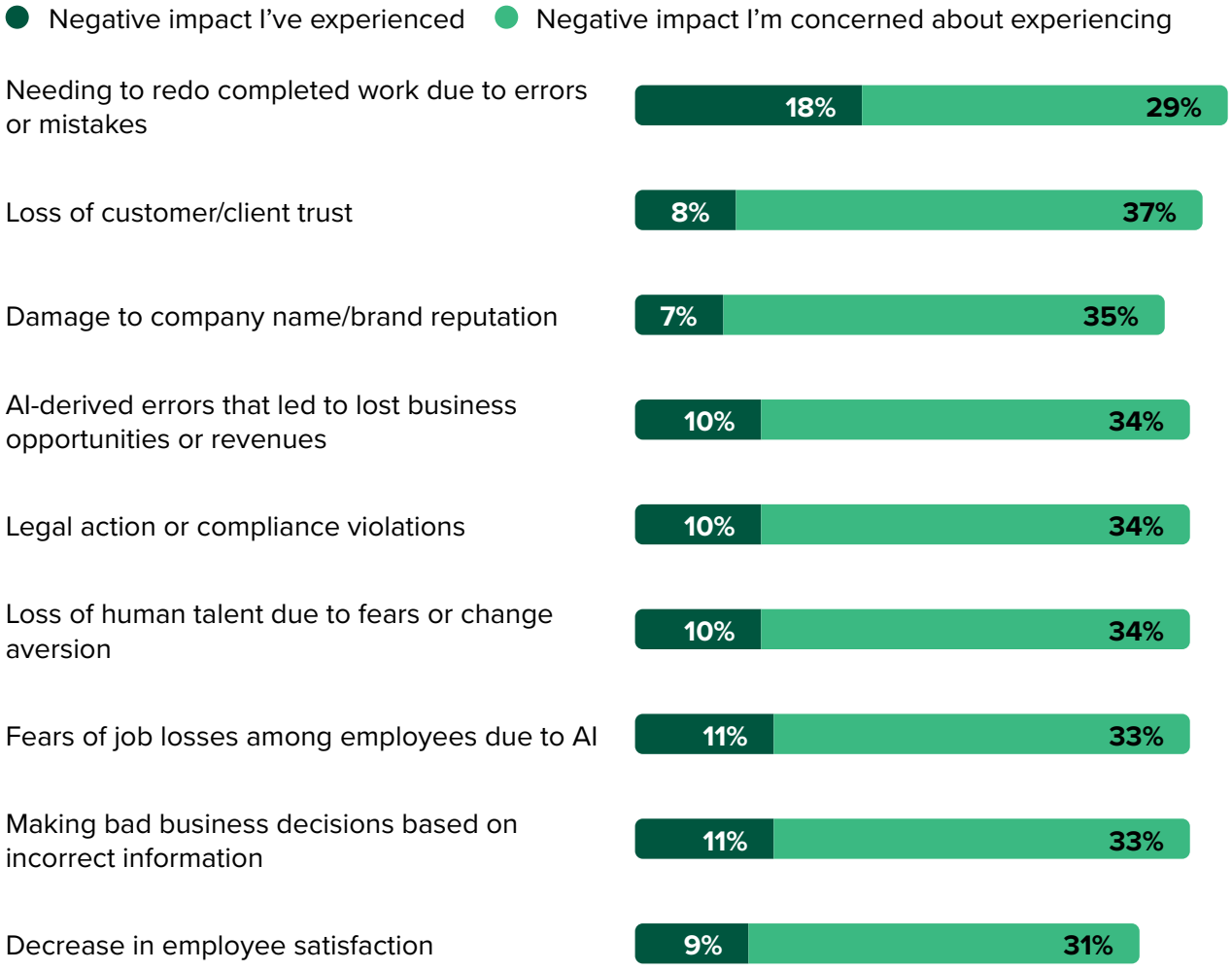
39%



Base: 608 global AI power users across various industries from a mix of practitioner and decision-maker roles
Source: Forrester's Q2 2026 AI Output Quality Survey [E-70447]

FIGURE 4

Potential Negative Impacts Of Poor AI Output Quality



Base: 608 global AI power users across various industries from a mix of practitioner and decision-maker roles
Source: Forrester's Q2 2026 AI Output Quality Survey [E-70447]

AI OUTPUT QUALITY: ACCURACY

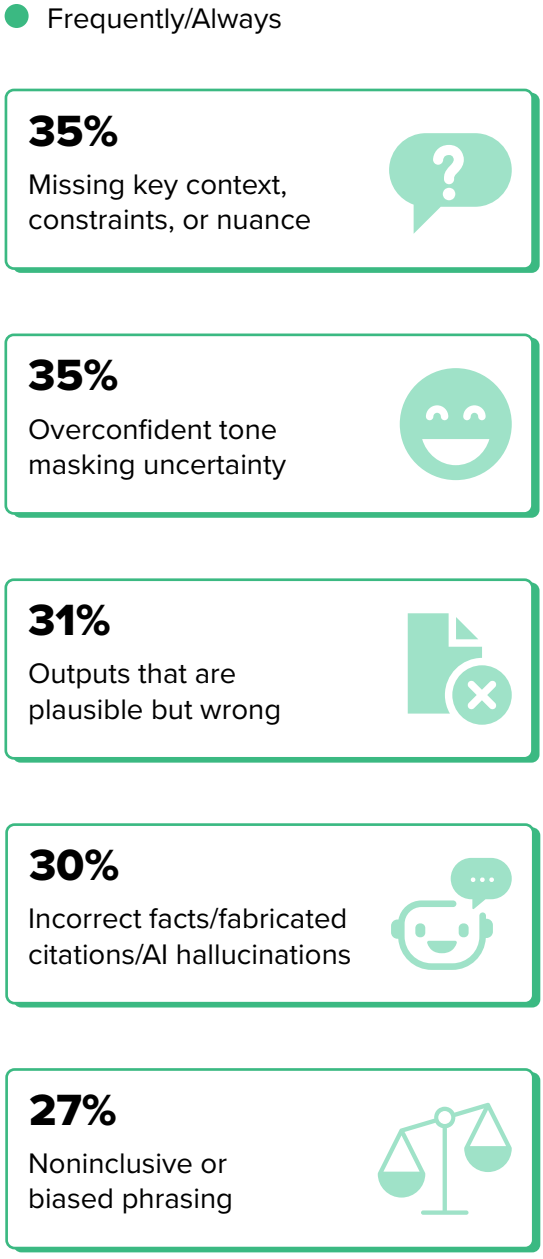
When asked to rank accuracy, speed, and trust, the majority of AI power users listed accuracy as the most important element of AI quality. This is likely because accuracy errors carry the heaviest potential consequences from the risk of outputs being used in ways that could negatively impact a business and/or its customers. Of more concern, however, is that 40% of AI power users said they only verify the accuracy of AI output "Occasionally" (or less often) before using it for work purposes, and only 17% said they

verify output “Every time.” It’s therefore unsurprising that only 32% reported finding errors that need correction “Most of the time” because with their inconsistent verification of information, error rates could be higher than reported. These numbers are alarming given these respondents’ frequent AI use.

The one-third of respondents who reported frequently encountering errors with AI outputs primarily found them in the form of missing key context or nuance, an overconfident tone, or incorrect facts and fabricated citations (see Figure 5). The frequency of these challenges among AI power users demonstrates that the need for output verification is constant. However, it isn’t practical, or even possible, to verify every single AI output across a larger enterprise. Knowing this limitation, the critical question is not whether to verify every AI output (because you can’t) but instead how to prioritize which outputs carry enough consequence to warrant verification efforts. Forrester’s AIQ data supports this, as a recent survey found that 51% of the general population of information workers don’t know when to question the output of AI, and could thus benefit from more structured guidelines for how and when to verify AI output.[3] This approach should allow organizations to properly design workflows that apply rigorous reviews to high-stakes outputs and accept residual risk in lower-stakes contexts.

FIGURE 5

“How frequently have you encountered the following with AI outputs?”



Note: Showing five responses
Base: 608 global AI power users across various industries from a mix of practitioner and decision-maker roles
Source: Forrester’s Q2 2026 AI Output Quality Survey [E-70447]

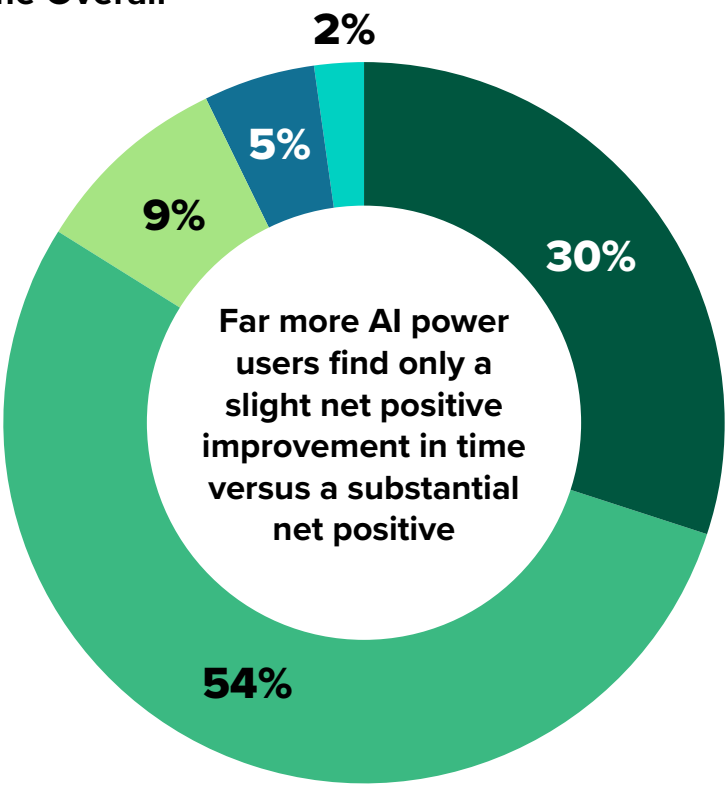
AI OUTPUT QUALITY: SPEED

Speed is another critical component of AI output quality. As a reminder, speed in this case is not about processing speed, but about the time it takes a human user to reach a suitable outcome. Overall, 80% of AI power users agree that AI is speeding up overall time to completion, however far more users find only a slight net positive improvement in time versus a substantial net positive (see Figure 6). The data illustrates this, as 74% reported typically needing several back-and-forth prompts before reaching desired outcomes, and 60% said they rarely get the outcome they want on the first try. Related challenges include incorrect or misleading information that requires extra validation (42%) and outputs that require too much editing to be useful (39%). AI rarely gets it right the first time, and the iterative process of prompting, verifying, adjusting, and reprompting are part of the equation for measuring the speed component of AI output quality.

FIGURE 6

Working With AI Saves Users Time Overall

- **Time substantially net positive:** Verification is very quick, which results in AI saving me significant time
- **Time slightly net positive:** Verification is fairly quick, which results in AI saving me some time
- **Time net neutral:** Verification effort is substantial enough that it negates any time savings from using AI
- **Time slightly net negative:** Verification effort is substantial enough that it takes more time to verify outputs
- **Time substantially net negative:** Verification effort is so involved that it takes significantly more time to verify outputs



Base: 608 global AI power users across various industries from a mix of practitioner and decision-maker roles
Source: Forrester's Q2 2026 AI Output Quality Survey [E-70447]

AI OUTPUT QUALITY: TRUST

The trust aspect of AI output quality is more passive than the other attributes. When employees are given permission to use AI tools at work, that permission implies that the tools (and the data) they use are safe and approved. The survey results validate this thinking, as 85% of respondents are confident that the way they use AI is safe and compliant with company policies. Only 27% (of the 85%) said they are “Very confident.” One reason for the lower percentage of high confidence may be that 46% reported oversharing at their organization, where AI made data available to employees who shouldn’t have had access to it.

Lack of trust in AI output can concern users enough that they will sometimes decide not to use AI for their desired task. However, this happens less than 50% of the time for most users, and not all tasks even have associated security or privacy risks. Even so, 25% of power users elect not to use AI due to security or privacy concerns more than half of the time. Two of the three most common reasons that these users choose not to use AI to complete a task center on trust: They are worried about data security (38%) or policy violations (32%). These findings are significant because even among highly experienced and informed AI users, trust is not a given.

Improve AI Output Quality With A Holistic Approach

Breaking down AI output quality into its separate attributes can make it easier to identify where to make improvements. We asked AI power users a series of questions about gaps that need to be addressed to improve AI output, investment tactics for AI tools, and general changes that would be most valuable in improving AI output quality. Across all these questions, the responses illustrated the need for better accuracy, speed, and trust:

- **For improving accuracy.** The goal of using AI to support work is to help create fast, quality output that can be used confidently, which requires accurate outputs. To bolster confidence in AI output accuracy, 45% of AI power users said they need to be able to trace how AI generates its results to help with validation, and 44% said AI needs a more accurate understanding of internal documents and data to drive improvement. Both of these changes would give users a better understanding of how AI derives its output, which is critical for verifying that results are real.

- **For improving speed.** Employees will get faster at using AI if they have a good understanding of how to engage with it. Thirty-two percent of respondents listed a lack of general AI training as a major limitation for enabling better output. Learning important skills such as prompt engineering and knowing when and how to verify references can have a huge impact on how long it takes users to achieve desired results. Due to this, 41% reported that their organization is investing in more training and learning initiatives to improve AI competency.

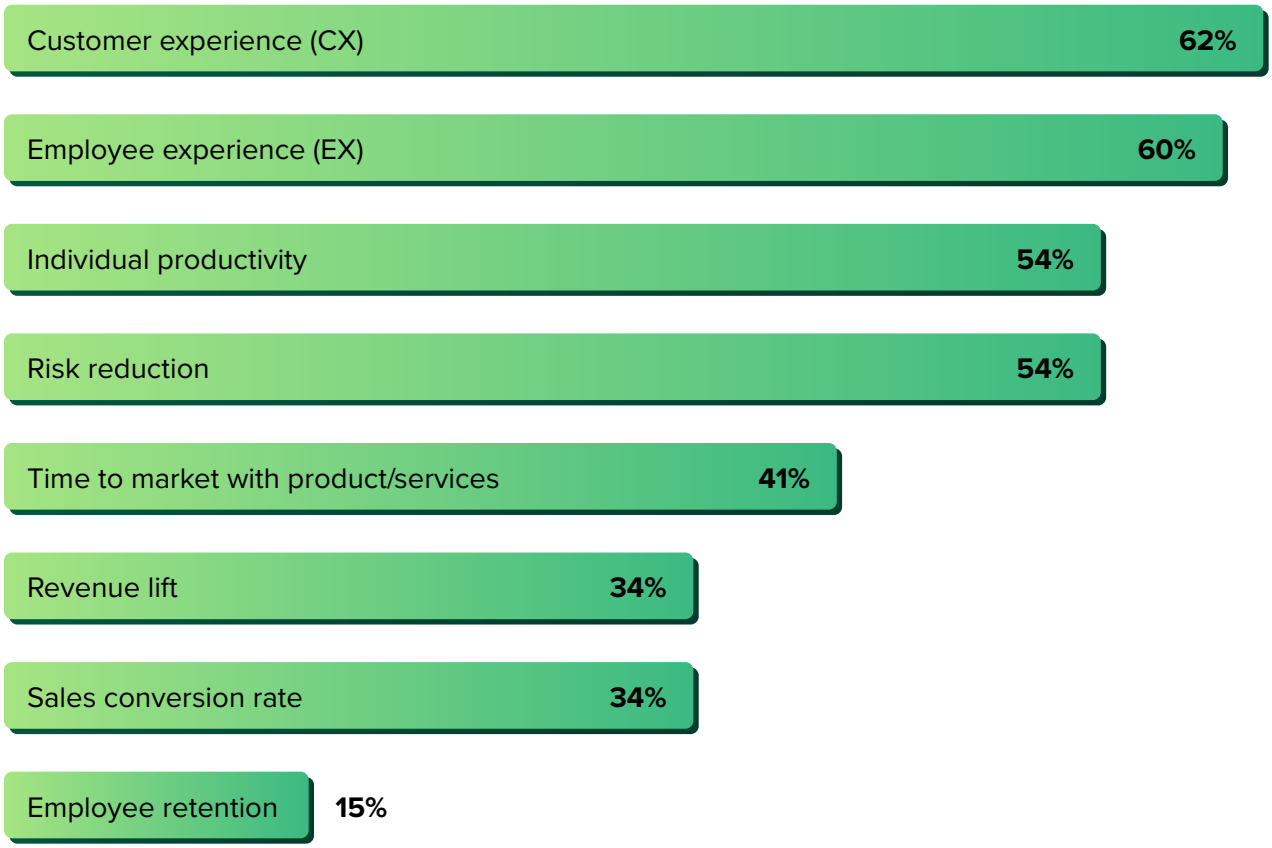
- **For improving trust.** More than one-third of respondents said that concerns about risk, privacy, and compliance is one of the biggest gaps they have in enabling better-quality AI output. If these concerns are left unresolved, the negative effect could impact perceptions of accuracy and likely reduce speed as users will be more prone to second-guessing AI outputs. That is why 38% are investing in refining and honing AI

governance and security policies to help employees be more confident in their use of AI.

No organization will succeed by taking a one-dimensional approach to AI output quality. By addressing the entire concept of quality (inclusive of accuracy, speed, and trust), organizational leaders will enable more confident AI use and be able to better track results. AI power users anticipate that AI output quality improvements will ultimately drive business improvements in customer experience, employee experience, and individual productivity and will reduce risk (see Figure 7).

FIGURE 7

"Which business metrics do you think will be most positively impacted by improving the quality of AI output?"



Base: 608 global AI power users across various industries from a mix of practitioner and decision-maker roles
Source: Forrester's Q2 2026 AI Output Quality Survey [E-70447]

Key Recommendations

AI power users in this study leverage AI deeply: 65% consider AI essential to their day-to-day job functions and over 50% use AI for more than half of their daily tasks. Even so, 49% harbor concerns about the quality of AI output. As a leader, you can improve your organization's chances of driving quality AI outputs through several key practices:

Raise employees' AIQ to set a baseline of human performance. While this study focuses on AI power users who possess strong AIQ, most white-collar information workers do not: Across five countries, only 16% score High AIQ.[4] AIQ includes skills like knowing when to question the outputs of AI, being capable of prompt engineering, and understanding how to use AI ethically. Lower AIQ compounds quality issues as employees do not apply the same level of scrutiny that AI power users do. So invest in raising the AIQ of all employees who will use AI.

Improve verification procedures with a human in the loop. This study found that too few AI power users actually verify the accuracy of the AI outputs they generate. Building systems that ensure the editing of AI outputs is crucial: Just as a journalist has an editor who oversees and edits their articles before publishing, AI needs human oversight before its outputs go out into the world. Build systems with human verification as an expected part of AI-infused workflows.

Ground AI use in rich contextual data. Accuracy, speed, and trust all increase when AI is grounded in the right data. Direct access to rich data sources behind the firewall, mapped into a knowledge graph, improves the chances that any given prompt will yield a valuable, high-quality result. Ensure that your organization architects its AI deployments with data context in mind to raise overall quality.

Build agentic AI process solutions for workflows. AI agents increasingly engage in co-work alongside human employees, acting autonomously and taking on tasks that involve many steps. AI agents can also raise quality: By optimizing and testing an agent against a given task, you remove the variability of users' skills; for example, not every employee writes individual prompts but runs the standard agent. Agents can therefore drive both productivity and quality.

Appendix A: Methodology

In this study, Forrester conducted a global survey of 608 AI power users in practitioner and decision-maker roles across Australia, Canada, France, Germany, India, Japan, Singapore, the United States, and the United Kingdom. Respondents were sourced from all industries, with a quota for one-third from the financial services industry. Respondent were from companies with at least 500 employees. All surveyed respondents were required to be heavy AI users, using an AI tool at least two to three times per week, and the majority ended up being employees who use AI multiple times per day. Half of respondents were in practitioner roles, and the other half were business decision-makers at the manager level or above. Survey respondents were given a small incentive for their participation in the survey. The survey began in May 2026 and was completed in June 2026.

Appendix B: Demographics

REGION	
North America (US and Canada)	**34%**
Europe (UK, Germany, and France)	**33%**
Asia Pacific (Australia, India, Japan, and Singapore)	**33%**

JOB TITLE	
C-level executive	**12%**
VP/director	**31%**
Manager	**7%**
Full-time practitioner/project manager	**50%**

PRACTIONER ROLES (TOP FIVE)	
Finance	**20%**
HR	**18%**
IT	**15%**
Risk and compliance	**11%**
Sales	**10%**

COMPANY SIZE	
20,000 employes or more	**9%**
5,000 to 19,999 employees	**26%**
1,000 to 4,999 employees	**45%**
500 to 999 employees	**19%**

MANAGER AND DECISION-MAKER ROLES	
IT	**19%**
Finance	**19%**
Risk and compliance	**17%**
HR	**17%**
Sales/marketing	**15%**
Product or service line of business	**13%**

Note: Percentages may not total 100 due to rounding.

Appendix C: Endnotes

[1] Source: Accerate Your AI Voyage, Forrester Research, Inc., April 2, 2026.

[2] Source: AIQ 2.0: Employees (Still) Aren't Ready To Succeed With Workforce AI, Forrester Research, Inc., March 23, 2026.

[3] Ibid.

[4] Ibid.


FORRESTER®