2025

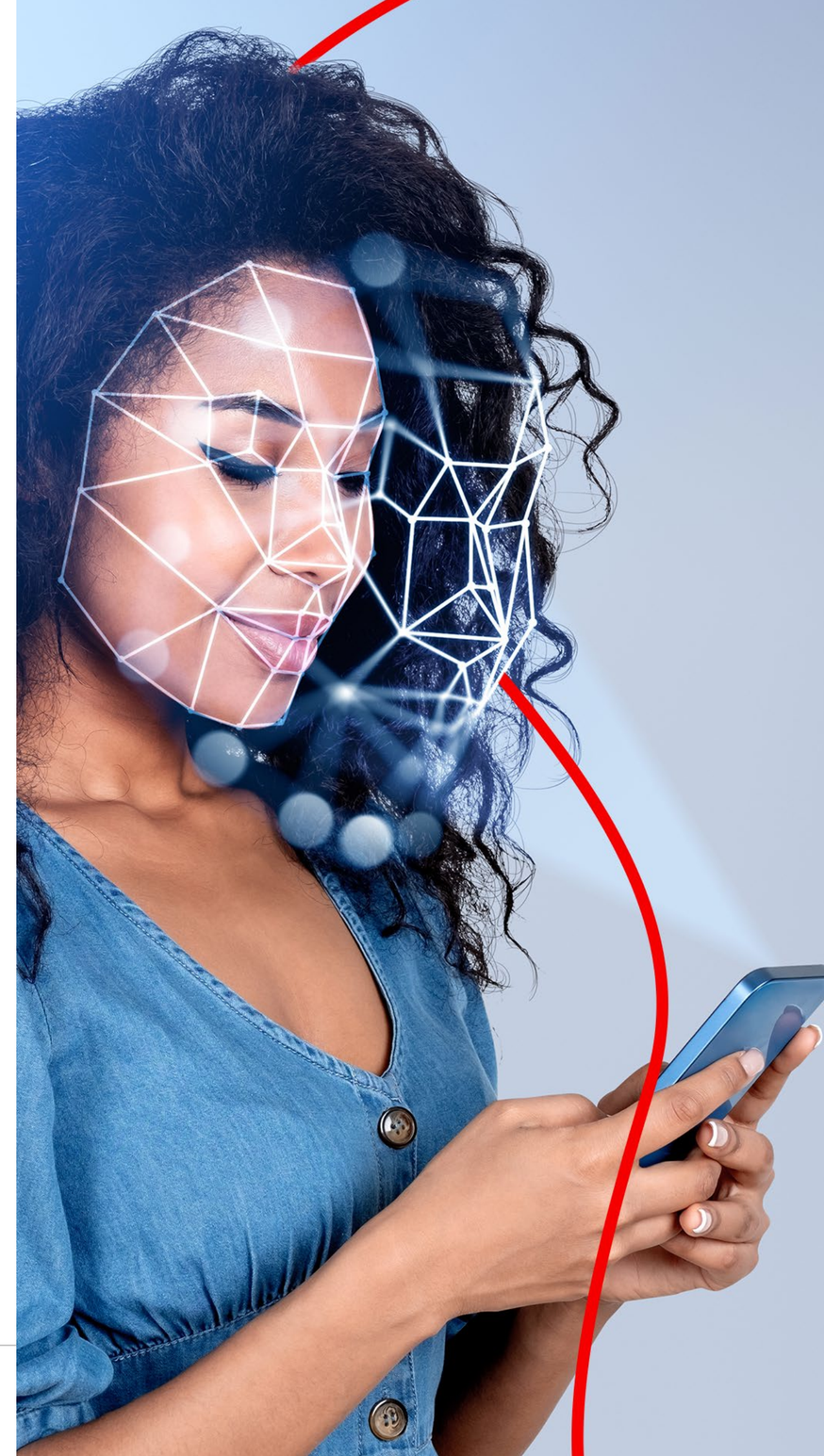
State of Digital Quality in AI

A snapshot of current trends, where AI fits in software development and testing, and how to create safer, more seamless AI experiences.

APPLAUSE

AI is changing faster than any other technology in recent memory, reshaping business operations and customer interactions across almost every industry. Software development and testing teams have a front row seat — not only are they creating and refining AI-powered experiences, they're increasingly incorporating AI into their own workflows. Applause surveyed software developers, QA professionals and consumers to learn how they are using AI, how use cases are evolving, where organizations are investing in AI, and how well AI experiences live up to user expectations.

In addition to survey findings, this report draws on Applause's experience serving global enterprises and technology leaders for more than 15 years, including many AI innovators. The report provides guidance on how organizations investing in the technology can gain the most value, whether they're creating AI-powered experiences or incorporating it into development and testing processes.



Executive summary

The State of Digital Quality Report in AI 2025 offers a snapshot of current trends in AI paired with key considerations for an effective strategy around developing and testing AI (as well as using AI in development and testing). Read on for this year's highlights.

- **Though AI is quickly changing how teams build and test apps and digital experiences, many organizations have yet to embrace the technology and integrate it into the SDLC.** More than half of development and test teams stated that their IDE has embedded AI tools — but 23.3% said their IDE does not (another 4.9% have no IDE). Among more than 1,400 software and testing pros, 7.3% say they have never used Gen AI for some of the most common development and QA tasks, others report they have only tried it once or twice. With 24.9% estimating that Gen AI tools boost their productivity by 25–49% and another 26.8% suggesting their productivity increases by 50–74%, incorporating AI strategically into the software development process clearly has the potential to deliver powerful competitive advantages.
- **Businesses are investing heavily in AI to enhance customer experiences and reduce operational costs — but flaws still often reach users, limiting the ROI of these efforts.** Nearly two-thirds of consumers using Gen AI in 2025 reported they encountered some sort of issue with the experience. Bias, hallucinations, answers that lack sufficient detail and misunderstood prompts are most common problems. As the market continues to evolve with Gen AI and agentic AI gaining wider adoption, comprehensive testing and risk mitigation efforts must be part of the release process for any AI experience.
- **Humans remain essential in testing crucial components of AI experiences.** While AI is increasingly becoming a part of the testing process — especially when testing AI applications and features — AI and automation alone are not enough. Humans must be part of the process. Rely on humans for subjective tasks such as assessing user experience, rating prompt and response sets, and performing adversarial testing to create AI experiences that are more likely to effectively serve real users in real-world conditions.

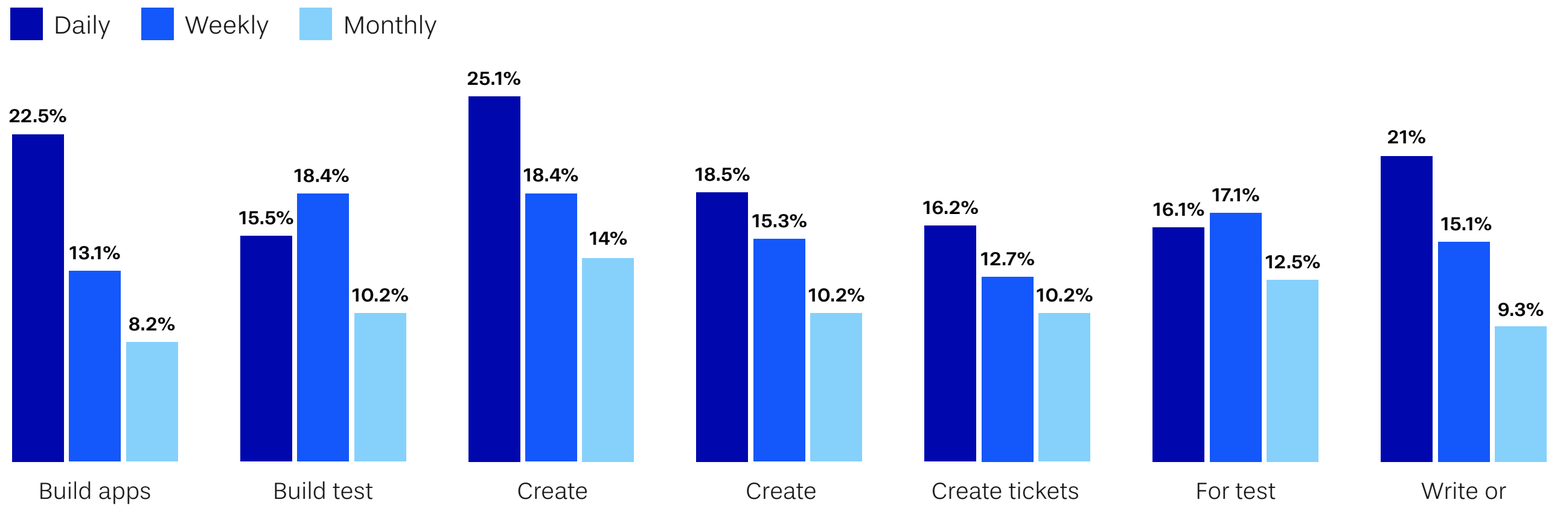
How Gen AI is changing software development

55.8%

of software development, product and testing professionals said their organization's IDE includes embedded AI tools. (n=1,644)

How often dev and QA pros use Gen AI for various tasks

n=1,575



Favorite Gen AI tools for coding assistance

n=1,070

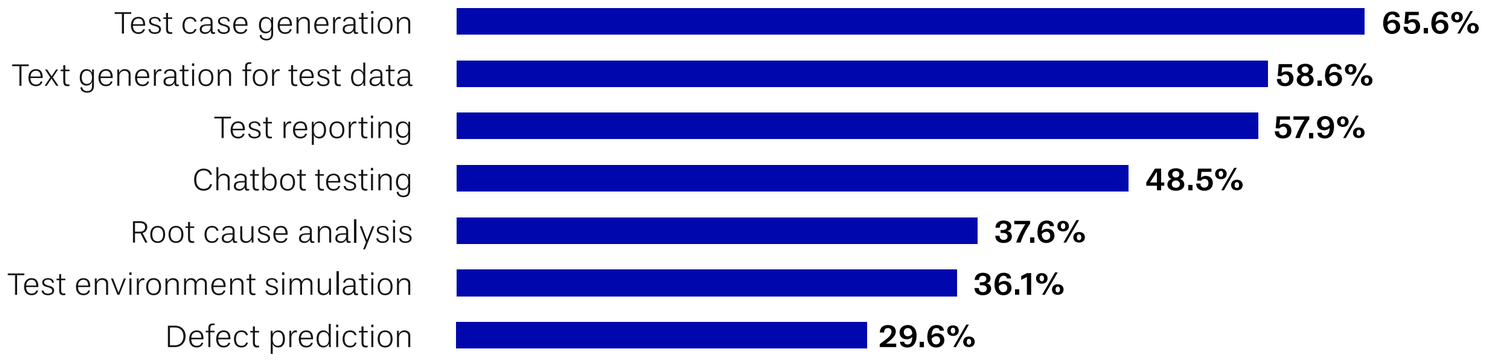
37.3%
Github Copilot

33.5%
Open AI Codex

7.5%
Deepcode

How teams are using Gen AI in the testing process

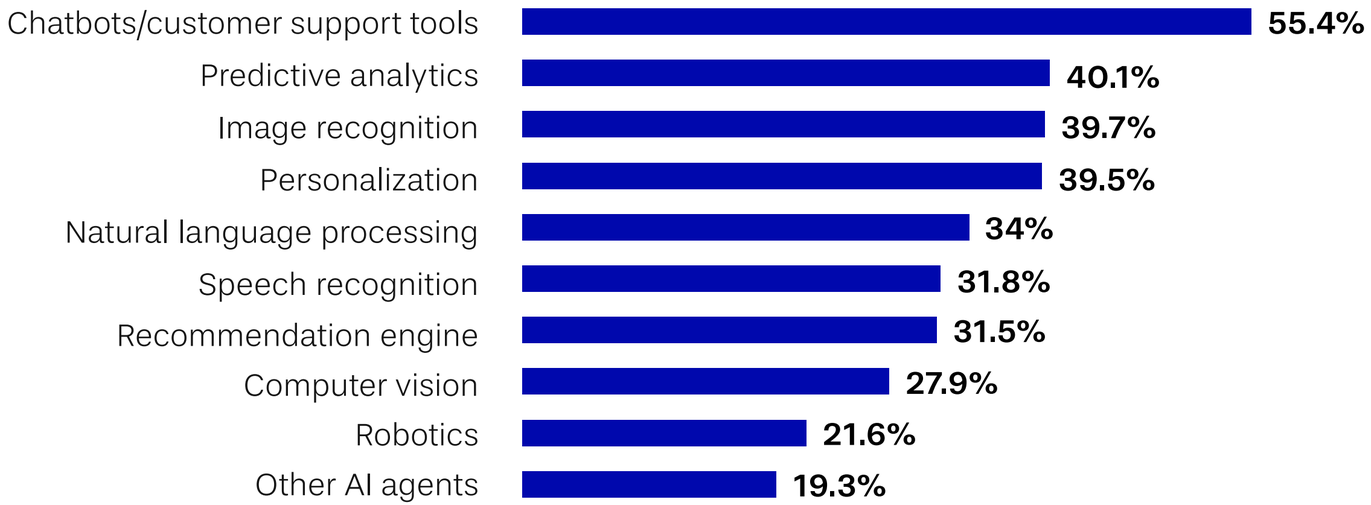
n=781



How teams are testing AI

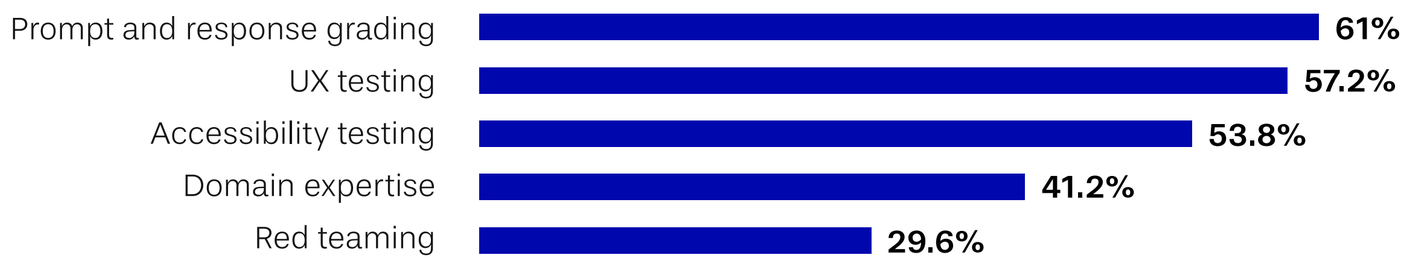
Types of AI features organizations are building

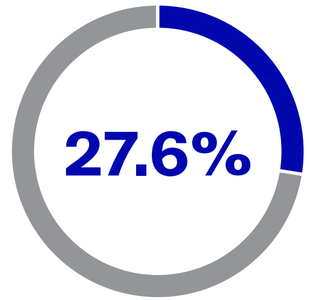
n=1,049



How humans are involved in AI testing efforts

n=621





Only 27.6% of respondents stated that their organization is not working on any AI applications or features. (n=1,592)

Techniques used to test AI-powered products or solutions

n=619

51.5%
Usability testing

47.7%
Black-box testing

41.5%
Metamorphic testing/ prompt and response grading

41.2%
White-box testing

38.6%
Explainability techniques

32.5%
Adversarial testing/ red teaming

Consumer experiences with Gen AI

29.8%

have swapped one Gen AI service for another and 34% prefer different services for different tasks (n=4,418)

77.5%

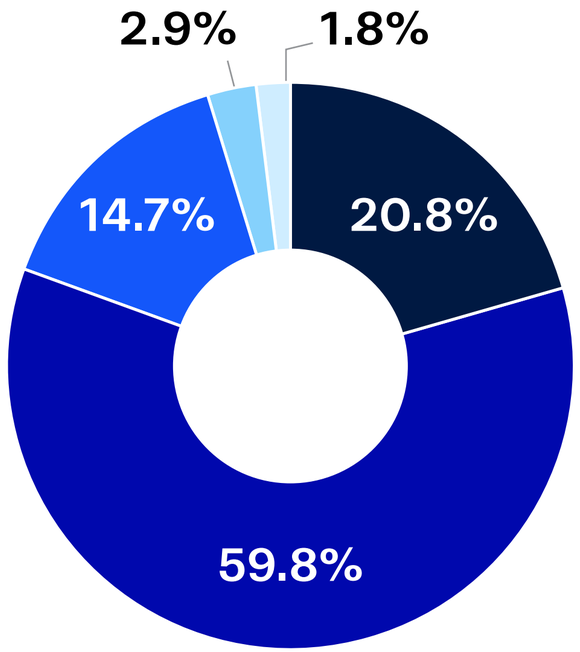
say multimodal functionality is important to them in a Gen AI tool (n=4,380)

34.5%

Only 34.5% report they have not encountered any problems using Gen AI since January 1 (n=4,185)

How often Gen AI understands questions and provides helpful answers

n=4,130



20.8% Every time

59.8% Most of the time

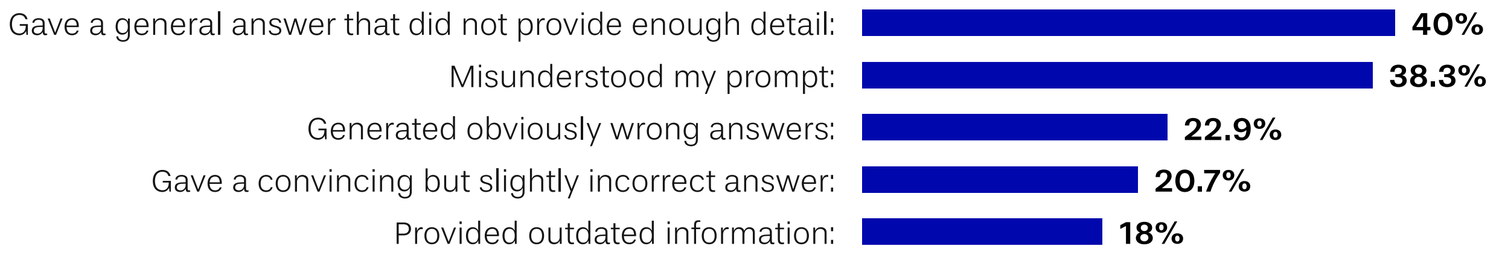
14.7% Sometimes

2.9% Rarely

1.8% Never

Most common problems encountered with Gen AI in 2025

n=4,185



Users reporting problematic responses in 2025

n=4,204

34.7%
Bias

16.9%
Offensive responses

32.3%
Hallucinations

Recommendations for delivering great AI experiences

1. **Embrace the value of real-world diversity**
   When refining AI models, particularly LLMs, it's crucial to acknowledge that data sets alone can't replicate the richness of real-world interactions. While domain experts and functional users may be able to adequately test internal tools and features, tools intended for a broader user base need additional scrutiny. Consider incorporating feedback from diverse user groups to uncover unforeseen biases and usage patterns. This approach can help ensure your AI performs reliably across a spectrum of contexts. Platforms that facilitate access to varied user demographics — particularly fit-for-purpose datasets — can be invaluable in this process.

2. **Prioritize continuous, iterative testing**
   AI model tuning is not a one-time event. Model drift and evolving user expectations necessitate ongoing evaluation. Implement a continuous testing strategy that allows for rapid iteration and refinement. This will enable you to quickly address emerging issues and maintain model accuracy over time. Leveraging flexible testing solutions that can adapt to rapid development cycles is essential.

3. **Focus on the human element**
   While automated tools are beneficial, the human element remains vital in AI testing. Subjective evaluations and qualitative feedback can reveal nuances that algorithms may miss. Incorporating human insights into your testing process will help you build AI that is not only accurate, but also user-friendly and contextually appropriate. Engaging experienced testers and domain experts who can provide detailed feedback can significantly improve the quality of your AI models.

"Companies developing AI applications understand that AI and ML technology are inherently more complex than traditional digital experiences, and come with more risks. To create meaningful value from Gen AI, organizations must include additional measures at every stage of the SDLC – yet do so in a way that doesn't slow innovation or delay time to market. With Gen AI's rapid growth and the shift to agentic AI, which introduces even more risks, businesses must prioritize a comprehensive strategy."

Chris Sheehan
EVP High Tech & AI

To learn more about the methodology of the report, visit **applause.com/blog/2025-ai-digital-quality-report/**

To see how Applause can help your AI testing and training visit **applause.com/generative-ai-testing**

Acknowledgments

Thank you to the business leaders, community members and subject matter experts who shared their time, knowledge and insight to shape this report.

Report Authors
Jenn Waltner, David Carty, Paul Hoffman

Creative Design
Joe Stella, Karley Searles, Megan Gawlik, Samsudeen Sallah

APPLAUSE