

ETR INSIGHTS

Insights Summary

Thought Leadership Webinar: Gen AI Impact in the Workplace, Part 1

President and Chief Marketing Officer | Tabnine

July 2024

ETR Insights presents an educational webinar series with Peter Guagenti, President and Chief Marketing Officer of Tabnine. Part one of the series discusses the state of generative AI and its impact on organizational workflows, particularly in software development, and the shift in executive mindset towards embracing AI. Mr. Guagenti describes how to best leverage AI for code generation, documentation, and unit testing; the importance of prompt engineering and context; the evolving capabilities of AI assistants, agents, and future AI engineers; lingering cultural barriers to AI adoption; and the need for robust privacy and data protection. Part two of the webinar series features two PhDs: ETR's Sr. Analyst Daren Brabham and Tabnine's founder and CTO, Eran Yahav, discussing the more technical nature of generative AI in code generation and Tabnine's product architecture. That should not be missed.

Key Takeaways

- Rapid AI Adoption:** Generative AI and large language models are experiencing the fastest adoption rate of any technology in recent history, with organizations quickly moving from evaluation to production.
- Shift in Executive Mindset:** A significant change in executive attitudes towards AI is driving its rapid adoption, particularly as executives recognize the competitive advantages AI can offer.
- Impact on Software Development:** AI tools like Tabnine are delivering significant productivity gains in software development, with improvements of 20-25% in tasks such as code completion, unit testing, and documentation.
- AI as an Assistant, Not a Replacement:** Current AI tools act as "assistants" that enhance the work of software engineers, rather than replacing them. These tools help automate mundane tasks, allowing engineers to focus on more value-added activities.
- Challenges in AI Adoption:** Data privacy, security, and regulatory compliance are major concerns that continue to hinder broader AI adoption, alongside lingering skepticism about the quality of AI-generated code.
- Importance of Personalization and Privacy:** Tabnine emphasizes the importance of privacy and personalization in AI deployments, ensuring that AI tools are both secure and relevant to each organization's specific needs.
- Future of AI in Software:** AI is expected to evolve from "assistants" to more autonomous "agents" and eventually to fully autonomous "AI engineers." However, this transition will be gradual.
- Private AI Deployments:** The future of AI in organizations may lie in private deployments, where companies maintain control over their AI models and data, ensuring compliance with both internal and external regulations.
- Competitive Landscape:** Despite competition from major tech players, continued focus on privacy, personalization, and system integration provides a competitive edge in the AI market for independent software development tools.
- Need for Early Experimentation:** Organizations must experiment with AI tools early to avoid falling behind in the rapidly evolving technology landscape.

Webinar Highlights

Tabnine, which created the first AI code assistant in 2018, has seen growth to a million monthly active users driven by individual developers who recognize the technology's potential. C-suite executives are also now racing to adopt and benefit from AI; Tabnine is second only to Microsoft Copilot among those deploying automated code generation. *"ChatGPT changed everything,"* notes Guagenti, *"because all of a sudden it demonstrated to us what was actually possible with these large language models."* Compared to customer support and chatbots, AI-generated software is relatively mature. *"We're still in V.1, but it's leapfrogging as we add capability, and I'm excited to be a part of it."*

Guagenti describes the greatest barrier to adoption as cultural; many still worry that AI will ultimately replace human developers. He encourages users to instead think of Tabnine as an *"Iron Man suit for the mind,"* making developers more efficient and effective. *"When we wrote applications before, we had to do things like memory management—and before that, we used to stuff index cards into a machine. We don't do those anymore, but there are still software engineers."*

Is your organization evaluating or has your organization evaluated Generative AI / Large Language Models (LLM) for business use cases? If so, for what business use cases?

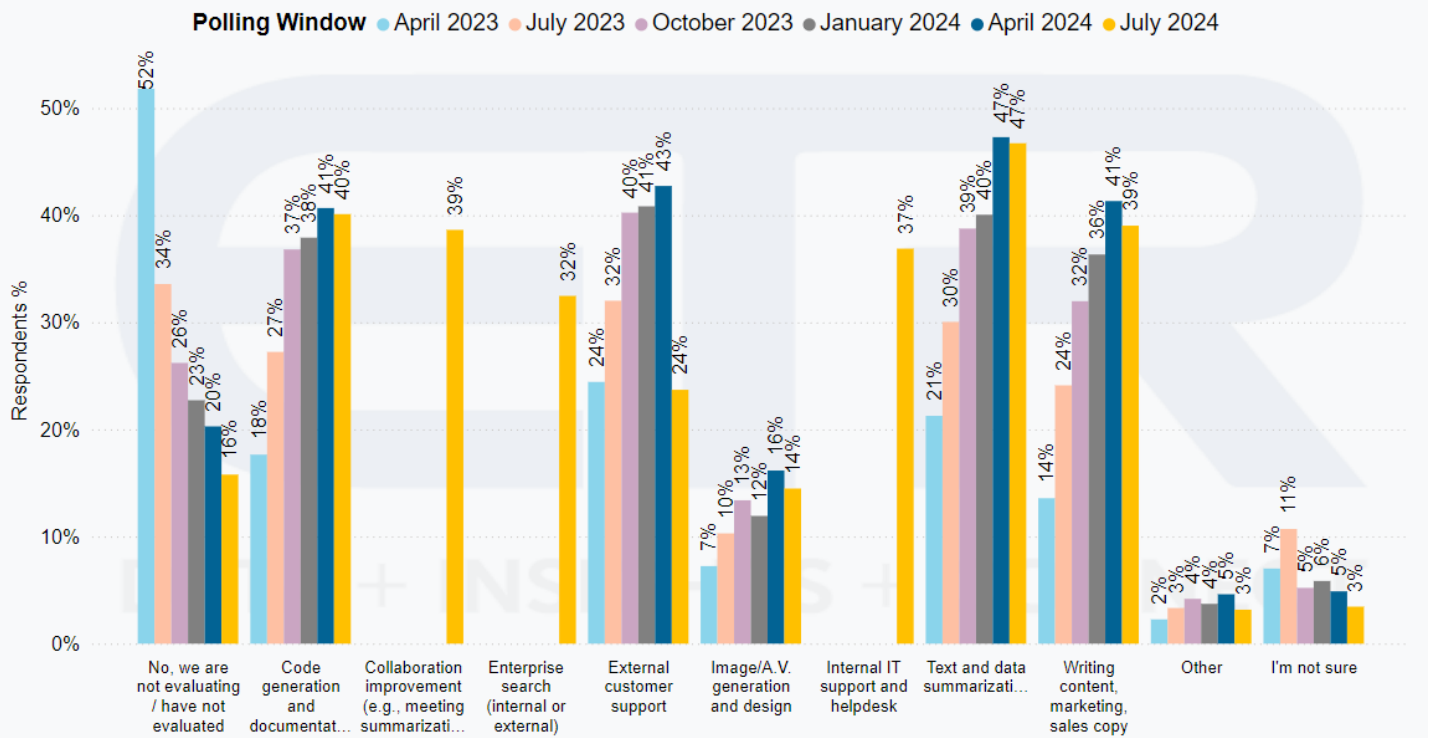


Figure 1. In ETR's most recent quarterly State of Gen AI survey, which was conducted in July of 2024 and captured responses from more than 1,700 IT decision makers, less than 20% of respondents say that their organization is NOT evaluating generative AI / LLMs. This number has consistently fallen each survey period. Text and data summarization, marketing and sales content writing, along with code generation and documentation, are the most evaluated business use cases.

Unit test generation and documentation are prime for automation; engineers will leverage AI to analyze how others have architected a problem quickly; AI onboarding agents will significantly reduce the time required for new engineers to familiarize themselves with projects. *"Today, we have AI 'assistants.' Think about them as accelerating and simplifying certain tasks. We're increasingly moving into what are called AI 'agents,' which are able to do things autonomously. You can leave them to their devices to accomplish certain tasks."* A recent Carnegie Mellon study demonstrated how AI software coaches helped junior programmers not only develop in an unfamiliar programming language but retain those skills. *"These tools were basically on-the-job coaches learning the task and the language, that stuck even when the tool was removed."*

Guagenti explains how AI has evolved from a "generalist" to a more expert-like assistant within a short period, now providing context-rich insights tailored to individual organizations. *"We use context to be able to provide explanations rooted in your company's standards, guidelines, code bases, and non-code information."* We are presently in an AI 'assistant' phase, wherein programming copilots suggest code to a human developer, but the industry is quickly moving towards AI 'agents' capable of autonomously iterating code. *"That's coming. I think the next 18 months, you're going to see AI agents replacing having to go and do these tasks manually."*

Ultimately, Guagenti imagines a future of automated software 'engineers' capable of generating applications directly from project specifications, or efficiently reusing existing proprietary code. *"We have a fully functioning AI engineer today that can take a Jira ticket and turn it into a simple node application."* Human oversight, however, will be required for the foreseeable future. *"When you start building a lot of these capabilities in, the complexity of software means the current model of just feeding something requirements and getting something back probably isn't going to work. It's going to be a different way of working with each other between the AI and the human, to do this where it actually works."*

AI-driven tools present a cost-effective alternative in a budget-constrained environment. To that end, leveraging existing, mature tools is typically more practical than building AI solutions from scratch. *"Would you go and build your own CRM system again? Would you go and build your own email platform again? You just wouldn't do these things."* Efficiency gained from such tools - particularly those that save engineers' time - means an almost assured ROI. *"Tabnine sells these things for tens of dollars a month. Based on the average salary of a software engineer, you need to save like an hour of their time in a month to pay for it."*

Guagenti explains that true accuracy hinges on relevance and context; he likens AI to an "ignorant genius," capable of broad knowledge but often without contextual understanding. True "accuracy" in AI responses is less about output from broad LLMs, but answers relevant to specific user needs. *"The second area of this is actually the prompt engineering, and taking what is the user intent, running it through the system, and coming back with something that the user believes is appropriate."* Most critical, however, is highly relevant context. *"If you can feed between the prompt and the LLM the most appropriate context—and note, I didn't say lots of context, but the most appropriate context—we've seen a dramatic uplift in the acceptance rate of code."* Tabnine aims to make its generative AI as vital to engineering teams as an "Employee of the Month," integrating deeply with a company's workflow, leveraging global code bases, Git repositories, and non-code sources like Jira and Confluence, to deliver customized and relevant results. *"Maybe the next generation of what we're doing is things like reading your observability platforms, and it has access to your identity and access management system, so it knows who has permissions to do what and how."*

Why has your organization's evaluation of Generative AI for business use cases not resulted in usage in a production environment?

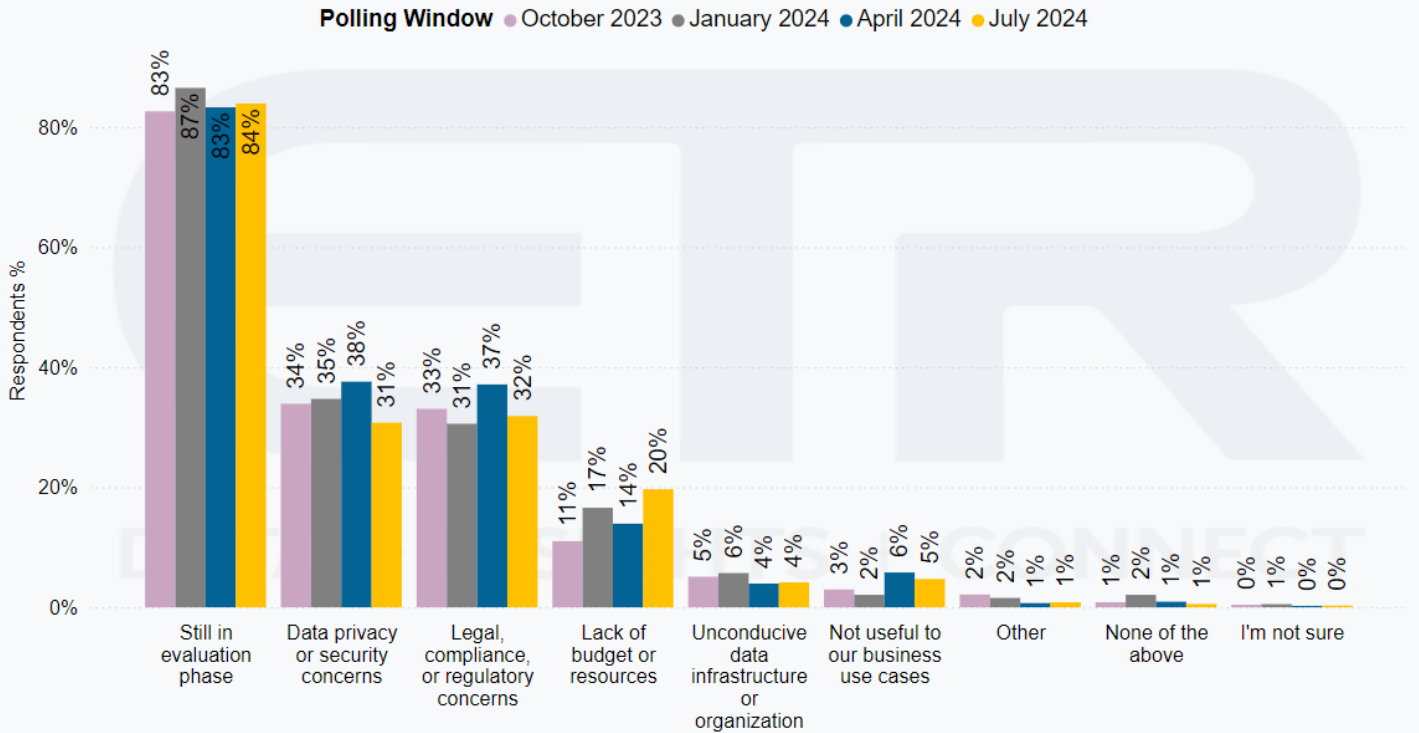


Figure 2. According to the recent State of Gen AI data, the clear leading reason why organizations have NOT moved generative AI / LLM evaluation into production is that they are stuck in the evaluation phase of the process, a number that is stagnant over surveys. Data privacy or security concerns and legal, compliance, or regulatory concerns remain the next most common reasons, respectively, softer survey-over-survey but still around one-third of respondents. Notably, the proportion of respondents citing a lack of budget or resources has risen to 20%.

To address customer concerns, Tabnine emphasizes three core principles—privacy, personalization, and protection. To mitigate risk, Tabnine has adopted a zero data retention policy, even in its SaaS platform, and all data is processed and wiped immediately after inference. *“Nothing is stored.”* Models are tailored to the specific needs and risk tolerance of their clients. *“They have our high-quality baseline models, we take their high-quality code to fine-tune it, and literally it’s a stepwise process they go through themselves, and it stays right there. We think that’s got to be a fundamental tenet of all of this going forward.”* From fine-tuned versions of Mistral to models trained exclusively on permissively licensed code and partnerships with Cohere and OCI that run on local infrastructure, Tabnine offers clients full control over the exposure and privacy of their own data, as well as the ability to switch between different models in real-time.

On competitors, *“Most people don’t understand what it takes to build one of these tools.”* Guagenti reflects on the intricate engineering behind Tabnine’s products; this accumulated knowledge, coupled with robust integrations across

major IDEs and Git systems, offers a formidable moat. *"We think there's more than enough of a market for us to be a strong independent business as we [serve] the 100 million people who touch software development in the world today, and we've built some really unique and killer IP to do that."*

The entire webinar replay and searchable transcript are also available from this webinar on ETR's research platform. In addition, part two of this educational webinar is also available, where ETR's Sr. Analyst Darren Brabham and Tabnine's founder and CTO, Eran Yahav, discuss the more technical nature of generative AI in code generation and Tabnine's product architecture. That should not be missed. Please reach out to us at service@etr.ai

Vendors Mentioned: Anthropic (Claude 3.5), Cohere (Command R), Google, Microsoft Copilot, Mistral, OpenAI (ChatGPT), Oracle, Tabnine

Contact the Insights Team to Discuss all the Details from this Webinar or Request Custom Research