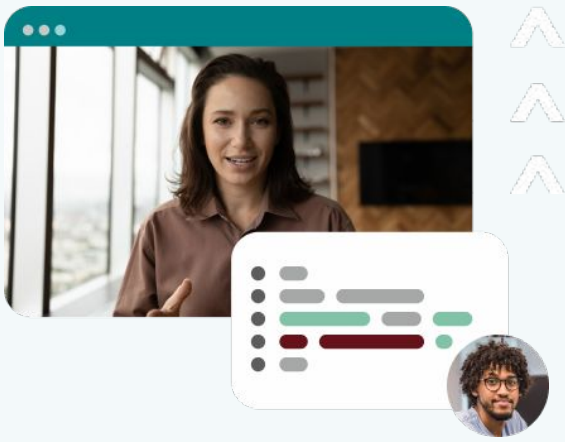
karat^

Interviewing in the Age of AI



Due to the buzz surrounding AI and speculation around the impact it will have on the way software engineers work, there are a lot of questions about the right way to hire in an AI-enabled world. Generative AI doesn't remove the need for programmers, but it can significantly reduce the cognitive burden of translating human ideas into machine code, greatly accelerating the ability to bring new products to market.

Most engineers aren't going to spend their days copying from LLMs, but just about every engineer is going to leverage them at some point. As such, AI is an additive skill set and should be viewed as a tool to augment and assist engineers in building the next generation of applications of digital experiences. To better understand the landscape of hiring alongside AI, we've tried answering some of the most prominent questions surrounding any uncertainties with AI in the hiring process.

In this document:

How are companies using AI today?

Most large companies are currently experimenting with some form of AI. This can range from using chatbots in services, idea generation, automating previously manual tasks, all the way to enabling AI-assisted coding tools.

Some tech companies are aggressively embracing the use of these tools and are starting to leverage ChatGPT or GitHub's CoPilot to get higher yields on productivity. New research suggests that more than 40% of software engineers are using AI coding tools on the job.

Many of the large enterprises we work with are also deep in exploration mode and identifying production use cases and are scaling internal teams. They are focusing heavily on leveraging AI to generate internal efficiencies (e.g. bill processing, call centers, etc). They are all pretty cautious, however, when it comes to customer facing products, especially in regulated industries (pharma, healthcare, financial services).

> E.g. one Fortune 100 CIO on our Technical Advisory Board stated **"I'm not putting AI into my pacemakers any time soon."**

Are companies using AI in their interviews?

While there has been a lot of interest in AI use during interviews, the sentiment around candidates using AI has been fairly neutral. Most of the questions we receive are about the impact of candidate AI use, and not specifically about enabling or disabling AI in interviews. While some organizations are experimenting with integrating LLMs into interviews, most enterprises are taking a more cautious approach. Very few large organizations are enabling AI in technical interviews today, although we expect this number to grow alongside enterprise adoption of AI tools in the workplace.

Note that this is separate but consistent with the approach organizations are taking with respect to Automated Employment Decision Tools (AEDTs). The current trend is for large organizations to exercise even more caution about using AI to make hiring decisions – in fact, none we know of do. This is due to several new laws coming out at state and local levels targeting the use of AEDTs. While we don't see much risk in our clients evaluating a candidate's ability to use LLMs in their work, there is risk if an org is using AI to score a candidate's performance and then using that score to make a hiring decision.

How does AI impact technical interviews?

The impact AI has on interviews differs based on the format and content. ChatGPT is very good at answering some questions and generating code samples. Asking someone to simply produce a working code sample and accepting their submission without human interactions invites the type of "cheating" that is concerning to hiring managers. This isn't a new problem, though. Candidates could previously break those interviews the same way by getting a friend to write their code. GenAI is just reliably available to everyone. Traditional code tests are particularly vulnerable to this form of cheating, which is why complicated "cheat detection tools" are increasing in prevalence.

Live interviews are different. Based on our extensive experimentation, we have found that allowing candidates to use AI tools does not disrupt or alter the results of well-designed live interviews with properly trained interviewers. For example, less skilled candidates who used GenAI to write code or solve a problem won't be able to discuss the approach they took or why they made certain choices. This becomes clear in a live interview setting. Skilled candidates using GenAI can write code faster than before, resulting in more time in the interview spent on problem-solving and less on producing the code. If this matches how they'll be doing the job, then it makes sense to interview them that way.

Does AI actually help candidates in live interviews?

While GenAI has the ability to solve some technical problems, it is not always correct, and often its explanations don't match its solutions.

Candidates that are overly-dependent on GenAI may be able to "trick" a code test, but the approach falls apart in a live interview. This is because when you ask today's LLMs a question, they have no notion of whether the answer is correct, only that the answer is plausible. During our interviews, ChatGPT and other systems frequently produced replies that were inconsistent – saying one thing while doing another, or claiming to fix an error while leaving the code entirely unchanged. LLMs have no sense that saying "the sky is orange" is false; only that it's much less likely to be an acceptable answer than "the sky is blue."

If a candidate asks how to reverse a string in Ruby, LLMs might provide a correct answer. But it might also grab the method that an engineer would use in Javascript and seamlessly adjust the syntax around it to look more Ruby-like. In our experimentation, our researchers have observed both of these scenarios happen. When a candidate is trying to answer a live interview question, it's less useful for them to trust a guess from ChatGPT, when they can probably find more definitive documentation by using a search engine.

Secondly, relying on AI to generate an approach and code is not an efficient way to solve problems in a timed interview. Writing out your question in a way that LLMs can understand isn't instant, and in our experiments, the turnaround time for responses was usually longer than a web search. Researchers also frequently needed to refine their questions with one or two exchanges, which added more time. If your question isn't answerable by a conventional web search, that delay might be acceptable, but in the vast majority of cases, there are other tools for developers that are faster and more accurate.

Overall, if a live interview is conducted correctly, we don't think that the use of GenAI will change a decision about an engineer who doesn't have the problem-solving skills that the role requires, and in many cases, over reliance on AI assistance will do more harm than good.

What is Karat's stance on using AI in interviews?

Our default recommendation is to mirror real-life working conditions in interviews wherever possible. For instance, we allow candidates to look up documentation from StackOverflow and Google, just as they would on the job. We think of ChatGPT and other LLMs in the same vein. As more organizations begin using AI, we expect it will become the default to enable interviews with the tools of the trade.

For organizations that allow AI tools on the job, this means allowing candidates to consult AI tools during interviews, and we have recently added a ChatGPT integration option for these clients to provide the best candidate experience. This doesn't degrade the hiring signal because candidates who are overly reliant on AI for coding struggle with code reviews and debugging portions of their interviews and in explaining why they have made choices. While GenAI can mimic a deep knowledge set and emulate comprehension, it doesn't do as well at analysis. This is an area we see strong candidates differentiate themselves whether they are using AI tools or not.

For organizations that do not utilize AI tools, this means turning off AI integration and monitoring the interviews for signs of external help. In these cases, our interviewers are trained to ask probing questions to determine what the candidate is doing, and they tag interviews and flag suspected cheating as needed for further review. This type of candidate behavior is rare, and the instances of it involving GenAI have been less frequent than candidates receiving help from 3rd parties off camera (which interviewees are also trained to detect).

How do you enable AI use in interviews?

The fundamentals of hiring talent alongside AI shouldn't change: hire people with strong problem-solving skills, and give them enough context and autonomy to solve your business problems.

Having strong interview content that is researched, tested, and continuously updated based on the latest developments is one way to ensure you're getting a true hiring signal on a candidate's abilities. ChatGPT is really good at taking a natural language description of a problem and turning it into code. And if all you're asking someone to do is translate from English to Python or JavaScript or C, ChatGPT can be really good at that, which underscores the vulnerability of code tests in an AI world.

The interview content and types of questions you're putting in front of candidates needs to evolve. You should design your interview format and content with GenAI in mind. Put your questions through a ChatGPT session and see how it does. Questions that are trivially solvable by GenAI are unlikely to be useful at discerning strong candidates from weak ones. If that's what you're seeing in your process, try to focus more on problems that are about logic and critical thinking where you generate a solution and then translate that solution into code. Another AI-resilient technique we use is to structure discussion questions around mini-code reviews. These questions require higher levels of cognition and produce highly reliable signals of a candidate's level of expertise.

To assess these fundamental skills, human interactions are crucial to truly understanding how candidates approach solving a problem and having strong interview content that is researched, tested, and continuously updated based on the latest developments is the best way to ensure you're getting a true hiring signal on a candidate's abilities.

There are three elements that allow companies to enable AI for technical interviews:

1. **Experienced human interviewers.** Coach interviewers to ask probing questions about the candidate's approach and problem solving based on their code. Observe how candidates test code as they work and debug errors. Focus on code review and other areas of analysis that allow candidates to demonstrate the expertise they'll need in an AI-enabled coding environment. These are all inherently human elements that will be absent from a candidate copying AI-generated code.

2. **AI-resilient interview content.** Interview formats that combine discussion with live coding create a robust profile of a candidate's abilities. This includes project discussions, code review, debugging, and live coding questions that require users to create a mental model of a solution as questions increase in difficulty throughout the interview.

   One way to generate AI-resilient content is to design interviews that require completion of a multi-step task. Working with an LLM on a complex task requires breaking the task down into smaller components. This design will gauge a candidate's ability to explain the task well enough to break it down into pieces small enough to prompt AI tools to offer the right assistance, so there is little risk of an inexperienced candidate succeeding on the basis of AI help alone. Another example would be to incorporate code security questions as LLMs cannot currently generate code that is consistently secure. Here you can challenge candidates to leverage an LLM to produce code without security vulnerabilities. Software engineers who are familiar with secure coding practices will be able to successfully complete this task.

3. **Ongoing content testing and feedback loops.** Karat is continuously testing, updating, and sunsetting interview content. This includes aligning the performance of experienced engineers, early career engineers, students, and nontechnical interview candidates against our benchmarks to understand how coders of different skill levels should perform. We test with and without AI assistance to ensure that our results are consistently delivering a predictive hiring signal.

How do you effectively prevent the use of AI in interviews?

The simplest solution to monitoring for this type of cheating is inherent to live interviews: ask a candidate to write code to solve a problem, watch them do it in realtime, and ask them to explain their solution. Observing a candidate writing code makes it obvious if a candidate is using assistance. A candidate who uses AI to generate a coding solution behaves like someone copying text from one window to another. The code is solved top down and in linear fashion. That is not how human engineers program.

In an experiment, our AI research team asked software engineers to guess whether a human or an LLM produced a code snippet based on a static image of the complete solution. Humans were no better than flipping a coin at determining which code-test response came from an LLM (and in one test, 75% of respondents answered incorrectly). Meanwhile, when provided with an animation of the code snippets being written, respondents were able to correctly identify the human 100% of the time. This is because the videos showed the thought process of engineers solving the problems inside out, manipulating variables, and testing at regular intervals.

Taking a human + technology approach to live interviews, and coaching interviewers to recognize the signs of inorganic behavior is a simple and effective way to differentiate between a candidate who has memorized or copied an answer vs. one who is solving a problem real-time.

There are many examples of organizations who do not allow AI use in interviews, including at AI-forward companies.

Glean, an AI-based enterprise search and knowledge discovery company, is one such organization. While acknowledging the utility of AI tools like Copilot for day-to-day tasks, Glean discourages the use of AI during interviews. Instead, the company focuses on assessing candidates' core skill sets without relying on AI. This is done through human interactions, rather than any kind of experimental AI-detection software.

> **"Ultimately that process isn't as difficult as it sounds,"**

shared Eddie Zhou, founding engineer at Glean. "The reality is having a human probing how candidates are solving problems at any kind of depth will be very apparent. It's just like knowing if a candidate has seen a problem before and is just regurgitating an answer. We've been able to acknowledge it and continue hiring at a healthy pace."

Duolingo is another company actively using AI to deliver customer experiences that is taking a more cautious approach to interviews. Here, the lack of AI enablement is something that resonates with candidates as part of the company's strong engineering brand.

According to Jocelyn Lai, Global Head of Talent Acquisition at Duolingo, many of their top candidates are eager to showcase their skills independent of AI. As a result Duolingo hasn't seen widespread requests from candidates to use AI. Lai elaborates,

> **"my hypothesis is we are quite well known for hiring and interviewing people for their technical craft, so, people go into interviews to show off how good they are."**

Other AI resources

- Research: Can ChatGPT pass a technical interview?,
  By Jason Wodicka, Karat Principle Engineering Advocate
- LeadDev: What AI has to offer: using LLM tools in interviews,
  By Don Gannon-Jones, Karat VP of Interview Content
- Webinar: Harnessing the power of GenAI in tech hiring,
  featuring engineering leaders from Okta and Glean
- Webinar: Harnessing the power of AI and the future of tech hiring,
  featuring talent leaders from Duolingo, Tinder, and Labelbox
- Webinar: How to build engineering interviews in the age of AI,
  featuring Karat's Don Gannon-Jones and Jason Wodicka