Interview Transcript – Participant: [Anonymized]

Date of Interview: 04 February 2025

Mode of Interview: Online (Zoom)

Interviewer: Anonymized

Participant Role: Software Engineer

Duration: 1 hour

Consent Obtained: Yes

All personally identifiable information has been anonymized.

---

**Interviewer**: Hello

**Participant**: Hello!

**Interviewer**: Thank you for accepting the request to participate in this interview and for being willing to share your experiences. Today, we'll be discussing examples of code cloning, specifically comparing AI-generated code with human-written code. Before we begin, could you tell me how many years of software development experience you have?

**Participant**: I have 4 years of software development experience at \*\*\*\* company.

**Interviewer**: I assume you're familiar with AI platforms like ChatGPT, Gemini, Claude, and others.

**Participant**: Yes, I'm well aware of them.

**Interviewer**: Do you have any experience generating code using these platforms?

**Participant**: Yes, I’ve generated code using large language models (LLMs) several times. However, I’ve never used them to generate code for the company I work for. My usage has been strictly for personal projects, primarily with ChatGPT, not for my workplace codebase.

**Interviewer**: So just to clarify, you're saying you've never used LLMs to generate code for your company's codebase?

**Participant**: That’s correct.

**Interviewer**: Was there any particular reason why you chose not to use this advanced technology at your workplace?

**Participant**: Yes, the company I was working for had a strict policy against using code generated by LLMs. This was primarily due to the sensitive nature of our data, and we were required to comply with those guidelines.

**Interviewer**: Did your team ever consider using LLMs locally to mitigate data privacy concerns? Or do you know if anyone tried doing that?

**Participant**: At our company, I believe a few junior developers may have been using ChatGPT, though nothing was officially acknowledged. Our codebase was highly customized and built on the Spring Boot framework. We also had to manage complex integrations across multiple frameworks, which made it quite difficult to generate complete workflows using LLMs. Because of that, we never explicitly used them in our development process. However, for some smaller feature developments, we occasionally used LLMs as a supporting tool, but not in any formal or consistent way.

**Interviewer**: Would you say you've used LLMs to generate smaller code snippets that were later included in your project’s codebase, even if not for full feature development?

**Participant**: Yes, you could say that. As part of our project requirements, whenever we had to build a feature, the user would usually provide a specific set of needs. Before starting implementation, we would analyze how those requirements could be integrated into our existing codebase. Based on the scenario, we often researched different approaches to achieve the desired outcome. During this process, we occasionally used ChatGPT to generate small code snippets. So yes, we did use it to support our project work in that way.

**Interviewer**: Do you believe that it is also AI-generated code, even though that particular code snippet is an AI-generated code snippet for your project?

**Participant**: Not always. For small programs or simple snippets, I think the AI-generated code often ends up being quite similar to what a human would write, maybe with slight differences, like variable names. So in those cases, it doesn't always feel distinctly AI-generated.

**Interviewer**: Okay, let's talk about code clones. Are you familiar with what code clones are?

**Participant**: Yes, I am familiar, and I know the types of clones. (also explained the defination)

**Interviewer**: (Gave some initial insights about how many developers are using AI-assisted tools for their day-to-day development process. Talked about some survey results by github and Googles analysis) We will go to the demonstration phase:

**Demonstration Phase – Summary**

During this phase, the interviewer presented participants with several code snippets: one set written by humans (sourced from Stack Overflow) and another set generated by AI tools, both designed to perform the same functionality. Participants were asked to review the snippets and identify which ones were AI-generated and which were human-written.

**Key Insights:**

* The participant couldn’t determine all the AI-generated code. (after removal of comments)
* The participant found it challenging to differentiate between AI-generated and human-written code, especially for common tasks.
* The participant observed that for smaller or utility-level functions, AI-generated code often resembles human-written code in structure and logic.
* The participant showed more trust toward human-written code when unsure, especially when the logic looked "too perfect" or generic in AI-generated snippets. This suggests developers may still prefer code with a known or trusted source, especially in production environments.
* AI-generated code was seen to follow clean conventions and good practices, which made it harder to distinguish from curated human examples. This may indicate that AI tools are maturing in terms of aligning with industry standards.
* Despite recognizing AI's strengths, the participant acknowledged that in highly complex systems, AI struggles to generate complete solutions.