# P2

Interviewer: ok, the first question, can you describe your main practices when you usually do code auditing? What is your entire process? For example, from how you initially download, and then understand the code, to finally finding vulnerabilities, and finally submitting the report. If it's ok, it would be better if you could describe this process with a detailed example. Also, could you share your screen and give an example? It'll be easier for us to discuss.

P2: Should I find a codebase to talk about?

Interviewer: Right, you don't need to show every detail, but you should show the key steps.

P2: Let me see if I can find something from my previous audits. I'll share it. [Screen Sharing] Can you see this in VSCode? For example, when we receive an audit project, after getting it, we first check how much code it has. This is a small project with only one file. If it has only one file, it has more than 300 lines, which is considered relatively small. If it's a small file, then I'll first check what it does. We can look at the name of the contract. It may also provide some technical documents. We'll check what the entire project is about, and then look at its imports. For example, here there's a 1155, and access control. Here it uses signature-related stuff. This is 12 lines. This is to prevent reentrancy, and then there's a library for string processing. Looking at the contract name, it's probably a game, a contract related to a game. This is just a guess. All these are guesses. Then at this point, we can use GPT. We can also ask GPT what this contract is about, what its functions are.

Interviewer: Let me interrupt. How do you ask ChatGPT when you copy code and then ask it in GPT?

P2: I'll just directly ask what functions this contract implements and what its purpose is, just a very simple question, and it will give a summary description, not too long. First, take a look at what the overall situation is like, and then ChatGPT will describe to you what kind of contract it is and what functions it implements. Next, you can ask about its global variables, for example, in this part from lines 28 to 52, what the purpose and function of each global variable is, and which business logic each corresponds to?

Sometimes I also ask it from a different perspective, for example, in which functions a certain variable appears, and which functions modify it. In this way, if I think this variable is relatively important and may cause some serious risks if not handled properly, then I will focus on this variable. I will go to ChatGPT and ask it which functions use this variable and which functions modify it, so that I can focus on those functions. After understanding some of the functions and roles of global variables, we can then enter each function. This is different from what I do in VSCODE. Generally, I will provide a piece of code to it and ask it, and it will describe to me that this code is for handling token exchanges, and it will give me a detailed step-by-step analysis of this code. Today, I tried a contract and directly asked it. For example, this structure is an auction-related structure. If it is an auction-related structure, I want to

know which functions it involves, and then I will ask it about the relevant business process. The sequence of the entire business process is the execution order of these functions. For example, it generally initializes first (i.e., init), and then the role and calling order of user-invoked functions.

Interviewer: Right, it's like when you ask it to explain, you'll specify which aspects of which code you want to understand, right? For example, variables, which functions, and the business process between different functions, how different functions are combined to complete a business? You focus on understanding these points to complete your understanding of this code.

P2: Since a contract has two trading methods, one is auction, which relies on the "auction listing" structure, and the other is fixed-price, where the price is fixed and not auction-based. There are two trading business processes within the same contract. If I need to distinguish them, I'll just let ChatGPT handle it, because if I were to check them one by one myself, I might mix up these one or two business processes. Then ChatGPT will tell me that the "init auction" function is for initialization, and users call the "byte" function to place bids, and finally call this function to complete the auction.

If the next price is fixed, it also initializes first, then calls this function to make a purchase, and finally ends the fixed price. The two are different, but if they are written in the same contract, it will be difficult to distinguish for those who are just starting out. So, I will ask about it by understanding the business logic related to this variable.

Interviewer: So different variables will have different associated business logic.

P2: Such a situation may exist, but the prerequisite is that you must first use your own experience or identify that it has two sets of logic.

Interviewer: So it's like when you're understanding this code, you first use your own judgment, then form a general internal understanding of this code, and then ask about the relatively minor points.

P2: This way is more efficient. If I can figure it out in a short time, it will be even more efficient. If not, I can also turn to ChatGPT and start asking from scratch.

Interviewer: Okay, in most cases, can you tell on your own or not?

P2: In most cases, people are just too lazy to read themselves and hand it over to GPT first.

Interviewer: But ChatGPT doesn't have the ability to digest the code of such a large project; instead, it can only handle a small snippet of code.

P2: Essentially, we identify them one by one as contract files; a project would be too large, so there's no other way.

Interviewer: Okay, if it's at the project level, you have to understand it on your own, right?

P2: Right, if it's a large-scale project, we can only first manually divide it into several modules, then further break it down into each module, and then use ChatGPT to learn and analyze such programs.

Interviewer: When you understand it, how do you make it help you?

P2: First, now it lists all these functions. For example, if I want to study this business process, I need to focus on these functions. If I want to find them through GPT, I feel it's a bit of a shot in the dark. For example, directly below, I'll ask, "Please analyze this function, which has some security issues. Please analyze the security issues therein and point them out one by one." When I input this paragraph, I'm not sure. I'm not even certain if it has any problems, and then I input it for it to point out. Then it will, based on its experience and potential security considerations, check whether the auction has started. It says that within the time limit of this function, it doesn't check whether the auction has started, and then it suggests making a clear check first. This kind of analysis is rather general; it's not tied to a specific problem in a specific piece of code. This is a bit of a shot in the dark, to see if what it gives is valuable, because if I have no direction, I'll just see what it says first.

Interviewer: Do you think that in most cases you first let it search and then it gives you some directions, or you have already come up with some directions and then let it confirm them for you?

P2: I think both situations exist. If I have no ideas, I'll teach it; if I do have ideas, for example, when there are some arithmetic operations within a function, especially those involving division, I'll tell it that division can lead to precision loss. Then I'll ask you to help me analyze the calculation issues within this function, and I'll emphasize the calculation issues, specifically whether there will be precision loss in the calculations within this function.

Interviewer: Let me check the question. You just mentioned different types of vulnerability. Okay, so when it comes to different types of vulnerability, do you have different debugging methods related to GDP?

P2: Specifically, under what scenario?

Interviewer: For example, some vulnerabilities are logical issues, while others may refer to variables, such as a variable not being allowed to be negative. When interacting with GPT, are there any differences in dealing with different types of such errors?

P2: I have tried to do this. I have used a numerical calculation myself, which is what I just mentioned. I pay special attention to the calculation issues of this numerical value, whether the parameter range is correct, whether it needs to be restricted, and then the issue of precision loss.

Interviewer: Did you write this prompt yourself? And it didn't provide this to you.

P2: No, this is based on my own experience. For example, regarding the problems that may arise in calculations, I just jotted down a few. Here, only these three paragraphs are written, and it is specifically used to analyze calculation problems.

Interviewer: Were these originally there or did you come up with them?

P2: This was generated by itself; I don't think I have any impression of writing these. Then here, it's my own note, and I just threw it up with a try-it-out attitude.

Interviewer: Do you think it's useful? Does it make use of this note?

P2: There is no particularly obvious feeling.

Interviewer: This note refers to yours.

P2: A case analysis of Hundred finance was attacked, and that attack was caused by the loss of computational precision, which is equivalent to a case. I feel that there are probably many pictures inside that cannot be uploaded, so if it can only view text, it cannot understand, because there are some screenshots that were directly pasted when doing the analysis, not in text form, and it can only upload text when uploading.

Interviewer: Okay, for example, if you've already found this error, what would you do? Would you verify it?

P2: Does verification refer to writing code?

Interviewer: The question is whether it is actually an error. That is, GPT doesn't just return a result to you, and then you have to confirm whether the result it returns has any errors. You may even need to write some attack scripts or something. So, how do you use GPT in this situation?

P2: At this point, it mostly relies on human effort, and it relies less on GPT. When it raises a question, it basically requires manual review.

Interviewer: Will you use it when you write the report?

P2: There is no report; everything is written manually.

Interviewer: In the current context of ChatGPT in the code auditing process, what challenges do you think there are?

P2: Is the challenge the inconvenient part?

Interviewer: Right, what are the areas that you find not user-friendly and hope to see improved?

P2: The inconvenient part is that I actually feel like I have to explain things to it slowly for it to understand what I want. From the very beginning when I input a contract, I have to tell it step by step to first analyze, and then guide it step by step to analyze that function and a certain variable. In fact, these tasks should be able to be handled by a dedicated GPT responsible for auditing in this area. As long as I provide the code, it will perform a set of standardized analysis and auditing tasks, without my having to execute them sentence by sentence and step by step all over again.

Interviewer: But have you ever tried, because I see you've written a rather comprehensive instruction here, and then have you ever tried directly feeding both the code and the instruction to GPT, letting it generate the desired results based on your comprehensive interaction?

P2: Never typed code.

Interviewer: You ask bit by bit, right?

P2: Right, because during the auditing process, I also got to know this project bit by bit. Of course, if it's like that, actually my expectation for its use is that it can assist me, but the ultimate result still depends on my understanding and knowledge of this project. I don't want it to be like a vulnerability scanner where I just throw the project at you, not even knowing what's inside the project, and then you just give me feedback on vulnerabilities. My usage scenario may be slightly different.

Interviewer: Do you think there are any challenges in finding errors?

P2: Yes, the biggest problem is that I feel a bit contaminated. I feel that when I ask GPT to analyze security issues, it always gives the same three things: integer overflow, reentrancy vulnerability, and another issue I can't remember. Anyway, as long as you ask it about security issues, it will report these, saying there's a reentrancy risk, an integer overflow risk, but these are a bit too general.

Interviewer: The scope is too broad.

P2: One issue is that the scope is too broad, and the other is that it doesn't conduct any analysis at all. It doesn't check whether there is a reentrancy vulnerability based on the code. It may just rely on some data or retrieved content, and when the reentrancy vulnerability appears frequently or in certain situations, it keeps reporting a reentrancy vulnerability repeatedly. Integer overflow and reentrancy vulnerability are two types that are often reported, regardless of whether the code actually has such issues. It will prioritize reporting these frequently occurring vulnerabilities.

Interviewer: Actually, you've answered this just now. When understanding this code, you felt that you needed to ask step by step before you could understand it, right? Are there any other challenges?

P2: Think about it again. If there are relatively many false positives, this is also a problem. However, there's nothing we can do about it. False positives in GPT are definitely inevitable because it's not a vulnerability scanning tool that makes judgments based on rules.

Interviewer: Do you think your efficiency has improved after using GPT?

P2: Of course there is improvement.

Interviewer: Approximately how many percentage points has it increased?

P2: Let me think. If GPT can get a good understanding of a project of any size in about a day, it would probably take me 3 to 4 days.

Interviewer: Look for errors. Do you think that, for example, within the same amount of time, the number of errors you've found has increased? Has your performance efficiency improved?

P2: Has efficiency increased? Yes, it has.

Interviewer: For example, you used to be able to find only one bug a day, but now you might be able to find a dozen or so in a day.

P2: Saying there are a dozen is an exaggeration, but there has been an increase. I've thought about some of the issues it pointed out. For some of the bugs it identified, the risk level should be relatively low. It tends to find more bugs with low risk levels or at the Information level, but for truly high-risk vulnerabilities, GPT can point out relatively few. While the quantity has increased, in terms of quality, the output of high-quality vulnerabilities is still relatively low.

Interviewer: You think so. We plan to design and develop an interface, software, or platform to help users better interact with GPT, to assist your work. Which stage do you think you most need help with? And what functions do you most expect?

P2: I hope it can provide more information in analyzing the overall project structure, especially if it can generate some business flowcharts or function call graphs.

Interviewer: It means all the business processes related to the list you just mentioned, anyway, just to help you understand.

P2: Right, it's about reverse-engineering from the code to the entire business process, understanding what kind of business it is, which functions users can call, and what role they play in the overall business. Because after we obtain the code and then look at its official documentation, the website's promotion, or its white paper, there are actually discrepancies, or rather, it's still not easy to understand. What's written on its official website sounds great, but in terms of its code implementation, it's hard to understand and difficult to correlate with what's on the official website. However, if you use GPT, it will explain based on the code what it's doing and what its business process is like.

Interviewer: During the process of assisting you in finding errors, what features do you most expect to have to better enable you to use GPT?

P2: What I most hope for is, of course, one-click generation of vulnerability reports. It should identify vulnerabilities and then generate vulnerability reports with just one click, just like a vulnerability scanner. This way, I don't have to check if there are any issues here and probe step by step. Just like a vulnerability scanner, it knows what it needs to do, then identifies, analyzes, and outputs the results in the form of a report.

Interviewer: Do you think you need to spend a lot of time verifying?

P2: What to verify? The result it outputs?

Interviewer: Right.

P2: It takes some time. If something is particularly outrageous, you can tell it's wrong at a glance. If there's something that seems a bit uncertain when it's mentioned, I usually write a demo to test it. But it rarely has the ability to raise a question that takes a long time to verify.

Additionally, I'm wondering what the business logic in a smart contract should be. Is it multiple entry points that then extend into different lines, with these different lines intersecting each other, but each entry point having a main line, which actually represents the so-called business logic of a certain business, such as an auction list? That is to say, within a contract, there may be many parallel businesses, and there may be intersections between these parallel businesses. These intersections may be function intersections, and they may also be intersections of state variables. I'm not sure if this description of the business logic of a contract or a business logic flowchart is accurate. I'm wondering if it's difficult for GPT to detect errors in such intersecting business processes.

Without providing any information to GPT, if we directly ask GPT to find a vulnerability, it is actually very difficult for it to discover these vulnerabilities through a highly reproducible process. It may discover them, but this process may not form a methodology, making it difficult to reproduce. However, I think that if, for example, my static engine or some other tools can provide such a trajectory, a main Line of Business or multiple parallel sub-Lines of Business, and can replace our current approach from function to scenario and from scenario to vulnerability in this way, I think it may be more effective. I think this is a very critical point

because this key point actually solves our current biggest problem, which is the context problem. The biggest drawback of our current use of a single function is that it cannot cover the context, and it is difficult to obtain the context. I think if this can be provided through some other means.

Interviewer: You mean going to draw that diagram, right?

P2: Right, I might, for example, it might be a point like this, then go to the next point, and then to the next point, you should be able to see it. This is another process, perhaps another process has reached this point. To put it simply, we actually want to extract different paths and different main business logics from the code. Then I'm thinking about this business logic line, but this business route is actually not very good. First of all, it cannot cover the business; it's just an incomprehensible fixed pattern. I think maybe we can use ChatGPT to extract this business-related process. The business process should actually be in our database, in our Knowledge Base, and there should be corresponding rules during the reasoning process.

P2: Just now when we were talking about business processes, I thought about it. This business process can actually be, I'm not sure if it can be regarded as a finite-state machine. For example, in a finite-state machine, it starts with a state, then you call a function, and it will enter the next state. How each state changes depends on the modifications made to those global variables. I think one thing that may involve human-machine interaction is that what has been presented like this in the past is actually relatively critical information.

Interviewer: How to make this information more effective?

P2: More effectively demonstrate it, whether using the text results returned by ChatGPT and adding some analysis of visual graphs, viewing visual graphs, or using other solutions to show the main business starting points, endpoints, and processes within the current contract, as well as the relationships between different business flows. I think this may be a relatively critical aspect in the audit process. Of course, this is also helpful for auditors. How auditors ultimately use this and how different auditors use it. I think that the creation and display of a process like the one mentioned earlier actually have an impact and are helpful for all three steps of planning, reasoning, and validating. Perhaps in the end, for example, the location where vulnerabilities occur may be in those relatively long processes or in those with more relationships with other processes. Maybe those with only one or two nodes and only one or two business flows may have a lower probability of generating vulnerabilities, which may be useful for planning.

Interviewer: So in fact, it's more of a process of how to display and consider visual aids. I feel like we've covered quite a lot today. Right, then let's stop here for now. Thank you.

P2: Okay, thank you.