P = Participant

I1 = Interviewer 1

I2 = Interviewer 2

I1

One thing in the beginning: did you notice any errors in the analysis session yesterday and the traceability or the evidence that we gave you?

P

Yes, I would say yes.

I1

So, the problem is that most of you found flaws in the model that we created. And the problem is that what we want to investigate here in this evaluation is the rule checking tool that works on the model. So, it checks the model and the flaw was in the model. So this is part of this evaluation. It’s a bit unfortunate. Maybe you can try to keep this in mind when we ask you about the tool that the input data was flawed basically.

P

All right.

I1

To begin, we would like to know a bit about your expertise. So, what's your job description?

P

I would say I'm a security researcher. So I'm doing in some cases security evaluations or working in EU projects like AssureMOSS. So mostly about security, IT security.

I1

All right, so how many years of security analysis do you have?

P

A little bit more than four years now.

I1

Four years, okay. And plain software development?

P

Before that, I was a software engineer, more than a... It was almost one and a half year. And before that, I was a university student.

I1

All right. So how many years of software development experience?

P

One and a half years.

I1

Okay.

P

One more thing: in Spring Boot and Spring Cloud I am not too... my knowledge is not too big in that, in this part.

I1

Yeah, no problem.

I2

What is the language? What is your language you're more familiar with? Programming language?

P

When I worked as a software engineer, I used C# and nowadays mostly Python.

I2

When you do your inspections, it's mostly Python projects.

P

Pardon?

I2

For the software, if you analyze and inspect, it's mostly Python.

P

Uh-huh.

I1

Okay.

P

Not always, but... it depends. If it's rather a backend then I would say Java or C# as well, in the front-end side, it can be anything.

I1

OK, then let's talk about the analysis session yesterday. So a first, simple question: when you compare the tasks where you did the manual part first and then had the tool output compared to the tasks where you had the tool output directly. Was there any... what was the difference in experience? Was it easier in one of the sessions, one of the tasks?

P

Yes, I would say when you have the tool output first, it gives you a good starting point. So from which you can - it's a good basis from where you can check other things, check the output of the tools. So you don't have to, I would say, fly alone, but you have some good points where you can start from.

I1

Okay. Maybe in comparison to this discovery phase that you had, is this something related? So when you say you don't fly alone, would this be something where the tool would be helpful as well, that you don't have to have this initial discovery phase?

P

So what do you mean initial? So without the initial phase or?

I1

Yes, I was wondering if you didn't have this initial discovery phase...?

P

Yeah, then it would be even more helpful because you have a lot of evidences, a lot of info about the whole code or the whole system and this can be a really, really helpful thing.

I1

Okay.

I2

So it's helpful but also from a conceptual complexity or effort, did you feel that it was making things easier or it was also, there was some complexity related to interfacing with the output of the tool or that was not a problem at all?

P

I would say the output was really convenient. So, it was really readable and you can check the evidence really easily. So this was really good. Well, I would say it helps, but if you don't know the whole system, it still requires you to check some things alone. So it's helping, but... but still you need to sometimes dig into the code.

I2

[Participant name], we should have said that at the beginning: we are not hoping that you tell us nice things. That's not the goal. So we are a researcher, so my life is a string of rejections. He's the younger one, so he's learning this, but I've learned it already. We have very big shoulders. We can take criticism and we want criticism. The goal is not to say everything is rosy and it works. We don't have a boss to report to that we did a great job. This is for science. If you tell us it sucks, it's good for us.

P

I would say definitely that I would check the tool output, but my impression was that it still requires some manual re-check as well. Because sometimes I feel that the evidences are good, but not gives the whole picture. So for example, in case of the entrypoints in the API Gateway, the evidence has that annotation about the code. But how I decided that it is an entry point is based on the Dockerfile that the port where the API Gateway works was exposed. So I would say it was a good evidence but not the whole picture.

I1

I see. If you compare the having it and not having the output, what exactly was the difference? How did you use the tool output?

P

The tool output makes it easier because I'm not really familiar with Spring Cloud. So this helps to... a little bit get to know things where they are and makes it easier to check some additional information I would say that I don't know to - on the internet that this is good, this can help, this is the good evidence. Without this it would be more complex to find these things.

I1

So was it mainly the linking to the code that you used?

P

Yes.

I1

OK. Just a very direct question, this step-by-step evidence -- going through this rule -- was this helpful at all, or was it mostly just an interface to the code, the places where you look at?

[THINKING]

I1

OK, it's a bit too specific…

P

Well, I use ... the output was helpful because it was readable, but I always checked the code where it points to. So, I used that.

I1

Okay. The next question would be if you trust the output, basically you said that already, right? You always double checked?

P

Yeah.

I2

So we have a few questions on trust in general, the trustworthiness of the tool and what is your level of confidence, and how do you...

P

Maybe it's a little bit difficult because even if I am trusting the results, though that I'm not always aware that it is right. It makes it a little bit difficult for me to trust just the output of the tool because I am not aware that it is right or not. And it's not because of the tool, but I don't know this Spring Cloud that much. So that's why I have to double check the things as well. But I would say, if I would know the Spring Cloud and I'm checking the tool output, then maybe it's -- I would say “yes, this is the good evidence” and I don't have to re-check the code or anything else because I would know exactly that the tool output is right.

I2

Yes, in a way it was also supporting helping with knowledge gaps, but in the same time, because you're not an expert, you don't know exactly how to take it. That's a sort of chicken and egg problem.

P

Yeah, yes.

I2

But... you did two tasks starting with the tool. Sorry -- two times you used the tool at the beginning and two times you used the tool at the end. In this progression... for instance, the second time that you started from the tool, did you feel better about it or more confident or was it still not much experience to change your opinion about the output of the tool?

P

I would say it was better because I got to know the tool, how they classify things. I would say the services, these things, “internal service” and getting to know how the tool works a bit, it gives me that kind of... so it makes it easier and... what should I say? So, I was more sure about that the tool is working for me. I mean, it's helping for me.

I1

All right. So, were there specific things about the output that made you trust it more or that made you trust it less?

I2

Or things that you did no-, you were very scared about?

P

For me, it was a little bit interesting that in the -- maybe it was the task two, the communication between the internal services -- I was not sure about why, for example, API Gateway communication between the “task webservice”, or I don't know what was the name of it -- is why not an internal communication. And I would say that it's still an internal communication. So they also should have been authorized. But as I remember, the tool classified it differently because API Gateway was infrastructural. And I'm not sure why it is.

I1

Yeah, I see.

P

And for example, also the database communication. I would say this also should be.

I1

Okay, yeah, good to know that it's more on the model side.

I2

Yeah, but maybe it seems to be more.

I1

So just as explanation for you, we get from this that we would have to explain these things more. So it's a conceptualization issue, I think, if you knew what internal exactly means in our context, that would have helped, probably.

I2

But this is also... what I gather... I don't want to put answers in your mouth, but I want to understand if I grasp what you said. Many things that you mentioned seem to me like that there should be a sort of engagement/agreement phase or protocol between you and the tool for what concerns the classification of certain concepts. What services there are and how they're labeled... and these are sort of fundamental information for the rest of the analysis.

P

Yes, yes. I think that would help because maybe this... well, maybe the problem was a little bit that I also was not aware about the tool. So I was... how from the DFD, the tool output was generated and what everything meant in that. And this makes a little bit also to feel that I have to double check the outputs, I mean the results, because I'm not sure how it was happening and what that really means.

I2

This is a very interesting point, thank you.

I1

Yeah, so knowing the tool better would have helped in your trust, probably?

P

Yes, yes, of course.

I1

Do you have any other suggestions that would help you trust?

I2

Not only knowing the tool, but knowing also...

I1

Yeah, the model and the concepts

I2

... what the tool is talking about, sort of a common vocabulary and a common understanding on the concepts.

P

Yeah

I1

Okay. Yeah, that's good to know. Is there anything else that would have helped you or that was missing? To trust the tool.

P

Let me just check, but nothing comes to my mind right now.

I1

By the way, if you need a refresher, we can also just pull up the output again, if you want to have a look at it to...

P

No, I have it here, I have prepared.

I1

Nice, very good.

P

Yeah... I am not sure that, for example, there is a tool output, and for example, in the properties field, when you don't have a link but you have an annotation that it is in, that's how it is coming if you don't have an evidence for that.

I1

Could you repeat?

P

Yeah, so I have a... in the first tool output rule, for example. In the properties field there are tags which don't have links on them. I think that this is because there is no evidence for that. And how can I say something without an evidence.

I1

I see. Yeah. Okay, also "no" in answer to that. So some of them are heuristics, there's no code evidence for it.

P

So it also comes from rather the DFD model.

I1

Yes. So, maybe to make you understand a bit more, there's this "internal" and "infrastructural" and that's basically -- basically we define everything that is like an API gateway -config server, these kinds of things- are “infrastructural” and everything else -that is the business logic- that's “internal”. But that means that there's no real code evidence for it. But for you to understand, I think we know what to do with that, we need explanations for everything.

I2

Yeah

I1

So we talked about the traceability already. You used it, right?

P

Yes.

I1

The links to the code.

P

Yeah.

I1

Was it sufficient for you? Would you have required more traceability for single annotations?

P

I would say these were good. In some cases, I felt that not all of them were right. Maybe one of my answers, I guess I wrote about it. But in some cases, I feel that this is not the right evidence based on which we should decide this. But these were a good starting point to check the tool these things first, and... so these were good and I use them, yes.

I1

All right.

I2

Can I ask you, sorry, just-

P

Yeah.

I2

Diverting slightly, but I think it's a terminology. So, professors, they like terminology. For you, is it clear when you talk about evidence and traceability -- or when we talk about traceability is it clear to you what the difference is about them? When you talk about evidence, do you mean the information that is given by the tool in the table, or also you include the fact -- the links that are pointing to the code? Just how you use the term, there is no right and wrong. It's just an understanding.

P

Evidence is -I would say- the hard fact in the code. And the traceability is how I'm getting there, I would say.

I1

All right. So, if we look at just the tool output at the table, we don't go back to the code, so assume there are no links in there.

P

Okay.

I1

This listing step-by-step, listing each node, say if they're included or excluded -- this general format of presenting the tool output, was this useful for you? Did you understand the layout?

P

Yeah, yeah, yes. I would say I really liked it because it starts with the statement, with a part of the statement. So, I always know where I am and what I am looking. Of course, “included” and “excluded”, sometimes I asked “why?”, but it was clear what the tool is saying for me and so I like this output form.

I1

Okay. Would you have... did you miss anything? Would there be anything that would further help you?

P

No. During the evaluation, I didn't feel that I would need something more in this output form. So I think it's good.

I1

Good to know, perfect. Taking a step back from yesterday, a more general context: in your daily work --- first of all, is this... the tasks that you did yesterday, is this related to what you do in your daily job? Is this somehow related?

P

I would say yes, yes, yes.

I1

Okay, so you do some security analysis?

P

Yes.

I1

Do you think you would use something like this? Assuming we have a proper tool, it's not just this one HTML document, bit of nice interface and so on… but just this general concept of automating these rule decisions, would you think you would use this and it would be useful for you?

P

Yeah, I would say yes. I would definitely start with it to see what the tool gives an output. So I would use this, but I have the feeling that I still need to check that it is right or not. So I would start to see what it is saying because it's a good -- I would say gives a feeling about the code and the whole thing, but I would check why it is saying it as well.

I1

All right. So more of as an initial starting point.

P

Yes, yes.

I1

Okay. That's good to know. Are there -- if you think about the tasks that this would be more useful for... you already said it's more of this discovery, but the rules that you could check with it, is it more helpful for very simple statements or these more complex ones where you have a lot of nodes or edges, like the type of tasks that you could do with it. What do you think it's best for?

P

For these rules, I would say it's really good. And this also can help in an evaluation. I am not sure that in, for example, some complex, more complex vulnerabilities, I am not sure how a rule can be made based on which the tool can help me to find that, or to say that “yes, it has or not has in the system”, I mean that vulnerability. But for things like this, “this how the communication goes” and “is there any encryption used”, etc., etc. This can be very helpful, yes.

I1

All right. And if you remember from my presentation, you can also write your own rules with the language of the tool. What do you think about that? Would you use this?

I2

Would you see yourself doing that or you like more the out-of-the-box "I use what is there"?

P

Well I would say, though I didn't try it out, I would say I would try this and based on that I would decide this -- if I would say “it requires more time for me to generate a rule that can give elsewhere” or which I can work with or “this too much time and so I can make my own assessment in less time without this”. So I would try it out and... although I didn't use it, I cannot say that in the long term I would use or not.

I1

Yeah, yeah.

P

But of course I would try it out because if I see that there are things with which I can make it quicker, I would use definitely.

I2

So you did not use the rules -- you saw them, so it's a bit difficult to judge, I understand. But from looking at the rules and the language, what do you think? Is it complex, is it simple, or...?

P

It seems a little bit straightforward for me. So based on the rules, what I see, it's not -- I would say it would not require too much time to get to work with these rules or define your own rules.

I2

Is it too simple, maybe?

P

I would not say they are simple but straightforward. So I would say they are not difficult to start to work with, I mean, generating your own rules. I'm saying that based on those rules that I'm seeing.

I1

Okay. So, just for understanding: you say it as a positive thing that it's easy to…?

P

Yes.

I1

All right. We're basically done with our specific questions. Is there anything left? Anything we didn't touch on, any feedback you have? Just open question, anything you have to add?

P

I think ()? my mind. It was interesting to see how this works and try it out. So, but nothing comes to my mind right now.

I1

All right, fine.

I2

Thanks a lot for taking time off of your job to do this for ()?.

I1

Yes.

I2

Very much appreciated, and thank you for the insights.

P

Thank you as well. It was interesting to try it out. Thank you.

I1

Okay, have a good day. see you.

I2

Bye-bye.

P

You too, bye-bye.