Opening Statement

Interviewer:
My name is Imane Akhyar and I am a Master's student in Information Studies at the University of Amsterdam. This interview is part of my thesis research, which I am conducting at [Company name] on how artificial intelligence can be effectively integrated into cybersecurity risk management within financial institutions. The interview is scheduled for approximately 60 minutes but will last approximately 30-45 minute. Your answers will remain anonymous and you may skip a question or end the
interview. Do you have any questions before we begin?

Respondent:
No as far as I am concerned, no.

Section 1: Background of the interviewee
***Interviewer*:**
Then we can begin.

Q1.1: Can you introduce yourself and tell what your role is within the organization where you work?

Respondent:
Well I am [NAME], I formally work for two organizations at the moment. Or actually maybe even 3, but I am officially employed by [NAME COMPANY]. That's a partner organization of [NAME COMPANY]. Think about the ins and outs of the organization, but also decisions about what we're going to do with the application [NAME] and if you know it. But also how we're going to roll that out to partners.
In addition, I actually work one day a week sort of on secondment from [NAME COMPANY] at [NAME COMPANY] to help out at existing customers. And I also work a few days a week at my previous employer and I'm downsizing. There I work as VP of Technology & Security and my role is to be involved in actually the AI initiatives that we have there. Both of how can we use within the organization? What are the security implications of that? But also how can we integrate it into the product? And then what are the aspects there that need attention now?

Interviewer:
Okay great so what I understand from this is that you do have experience with artificial intelligence then and know what it is and entails.

Interviewer:
Okay great!

Q1.3: And to what extent are you familiar with LLMs or other AI technologies in your work context?

Respondent:
Yes well I I, I know how they work conceptually. I'm involved in the choices of at least at [NAME COMPANY], which one we use, which tool we use. I know the concepts like bias, LLM, machine learning, all those kinds of aspects. I don't have specific training in it, so all the nitty gritty details shouldn't ask me, but conceptually I know what they do and how they work.

Section 2: Thematic depth

Interviewer:
Yes great that was then the introduction questions the next part is then the core questions that out yes that has to do with the frame on that I sent yesterday. I don't know if you've gone through those yet but most of the questions are either a combination of the elements or questions of single elements and at the end of this section is a contests also briefly discuss.

Well, the first question is quite a long question, I also noticed in the previous interviews, so I just have to put it this way.

Q2.1: In what areas within cybersecurity risk management do you see the most potential for LLMs to add value?

***Respondent*:**
Such a long question and a broad question.
Just to think about it.
I think LLMs can help in several aspects, but one is just in the formulation of what control measures do you want to put in place and how do you connect those to practice on the one hand and also are they testable. A second aspect is, I think that LLMs can be very helpful in making it understandable to people who don't have a lot of knowledge of cybersecurity and or risk management. That they can explain that in an accessible way. Without having to put a lot of effort

i have to put in a lot of effort. Because you often see that professionals quickly talk in terms that are difficult for non-specialists to follow. I also think that LLM is ultimately based on configuration or evidence is much faster and more objective than people can judge whether something is basically configured securely. Yes, then to no and whether measures are effective yes, then to no. And fourth I also think that LLMs can also help in the detection of anomalies.

***Interviewer*:**

Yes, right, because I was also thinking so in the detection of abnormalities I think they are really super useful for that. But of course you also have different types of LLMs. **Q2.1a: What types of LLMs do you think are most suitable for application in your industry, and why?**

Because you have, for example, GPT-4, Claude, DeepSeek, Apple Intelligence and so many more.

Respondent:
I think that's the phenomenon of the times anyway that you have to make a choice now between which model you want to use. I think in the future that will convert back to each other and that actually you just ask a question and based on an initial assessment by a kind of mini LLM the provider itself decides which model it's going to send it to and you don't have to choose that as a user anymore. That's my expectation, but it depends on what you want to achieve, doesn't it? For example, for assessing is there currently an anomaly in patterns and detecting incidents. Would you actually want to use ***reasoning models*** and for that example, what I mentioned of explaining to users, you would want to use much more of a ***general purpose LLM*** for that, because you don't actually have to use reasoning there, but more simply. What I think the the advantage of Claude over ChatGPT is that Claude is much more context driven , so you can give a larger context with that, which for complex cybersecurity problems can certainly be an added value.
But that's also a bit of a rat race, because now Claude is again the one with the biggest context and that you have to wait again for ChatGPT to come up with a new model or an adaptation of an existing model that the context
the window goes up so.

***Respondent*:**

Look, for the Western world I would never look at DeepSeek. I actually think there are only two relevant providers right now: Anthropic and ChatGPT or actually openAI. And you can see that a lot of the other propositions are based on that or can't measure up to that. GenAi is pretty funny, but can't measure up to the latest models from Anthropic or from OpenAI.

***Interviewer*:**

And then a question about within your organization itself, because you just said you work at 3 organizations.

FQ2.1a Has the LLMS already been experimented with or deployed in certain applications or is that still a gray area within the organizations?

***Respondent*:**

Well at [NAME COMPANY] they are now encouraging you to experiment a little bit with the free version of ChatGPT . That's funny, but I think that's insufficient, I think actually you should always use the paid version or market where you could really get serious value out of it. And within [NAME COMPANY] there is a bit experimenting with Cursor AI. That's specifically a tool in the area of software development and AI assisted coding.

Which again the actually just offers the models of an anthropic, Google, and OpenAi behind it with their own specific training to be able to do AI assisted coding efficiently. So we're kind of experimenting with that. But obviously it's just a very small organization, so it's particularly living on what [NAME COMPANY] is doing. [NAME COMPANY] is really several steps ahead. They recently did a proof of concept with ChatGPT Enterprise with 25 employees. And there the conclusion is that that delivers about 10% efficiency on average per week per employee. And, of course, that's just super interesting, because even though the license costs something, 10% is really significant.

And, that's without people being trained on how to use it or whatever it is. Again, the question has already been asked there, can we get that percentage up even further if people start training? Well, I don't know, we're going to see that, but with that, it's actually also a no brainer to roll out Chat GPT for the entire organization.
On the other hand, we also made a call last year with including an AI feature in the product that [NAME COMPANY] provides or one of the products

that they provide and the PoC (proof of concept) was basically successful. Only then do you see that it is difficult to go from a working PoC to production-ready functionality. Not just technically, but certainly also in terms of compliance. Because data goes to external parties anyway, and since this also involves matching paid bills and incoming payments, you also have to deal with personal data there. And that's just an interesting case.

***Interviewer*:** Yes and [NAME COMPANY] Are then obviously a step further than [NAME COMPANY].

Q2.2: How are AI systems monitored in your organization?

Respondent:
No, there is still relatively little for that. It was what I said, really a proof of concept so There are a few basic rules of use your mind and things like that but the advantage of course of ChatGPT enterprises is that you don't have to think about what you're allowed to put in, because it's not used for training or whatever, because you're just really paying for it in Euros, so there also have to be a lot less ground rules about you're not allowed to throw customer data in there or you're not allowed to do this or you're not allowed to do that.
Where I think still, the challenge is in whatever the AI ACT calls it, AI literacy. I think there are too many users who actually have no idea how you use it and how it works.

***Interviewer*:**
So that the education of that is missing, so that kind of a knowledge session and education courses for that should give word.

Respondent:

Yes, exactly. That's kind of useful and actually necessary for any company.

***Interviewer*:**
Yes, you had just included compliance, but of course you also have other legislations like GDPR Dora and the AI ACT.

Q2.3: How do regulations such as GDPR, DORA, NIS2 or the AI Act affect the application of LLMs in cybersecurity?

***Respondent*:**
I think that depends on the use of AI.
What I said if you look at the use of AI in day-to-day work, you actually have relatively little to do DORA and NIS2, GDPR of course you do but no different than what already exists for other systems that were used. So that's also nothing new for most employees and they are already trained in it. The gap is in the AI Act, which is relatively new. Think a lot of organizations have yet to look at that or what does this mean specifically? What do I really need to do to comply with this? Those kinds of aspects, but also there I think it's not used in the workplace because a lot of people will be using ChatGPT as a search engine anyway.

Of course, we also discussed it in the AI session last week. If you're using Google now, you're also using AI in the background, so I don't see that being essentially different where compliance plays a very big role. So if you are going to offer that in your product and you are the one who buys the AI huh? Who says, has the contract with the LLM provider, because then suddenly DORA becomes incredibly relevant, because then it's a subcontractor. Then it's a potential sub-processor. All those kinds of aspects of compliance. And then it actually also suddenly becomes a lot more relevant, because you're also reselling what you're doing there to your customers and potentially there the customers of it are enjoying it again. If you don't do it quite right. So I think it's much more relevant there. I think there are organizations that are good at compliance and I think we are organizations that are not very good at compliance.

Interviewer:
That's clear, before that you had included something about little knowledge. Of course, that is a risk, which is common with AI.

Q2.4 Are there any other ethical risks you see with the deployment of LLMs?
So for example bias and privacy are kind of the main risks that people are concerned about. But do you personally think there are other risks as well?

Respondent:
Look obviously what you always see is the more you automate in an application and I do see what AI and a potential to add another level of automation. That also means that there's just the headcount that has to go down. Also in that example, what I gave of that 10% efficiency at [NAME COMPANY].

Do CEOs say yes very nice, but if we're going to roll this out, we also just want to see a target per department of how many staff go out or
how many vacancies are you not going to fill Because you're getting AI now. That's not so much a consequence purely within the LLM itself, but rather of AI in general. An ethical piece of that is to what extent should technology replace people's jobs? And for some jobs I think that's very good and for others maybe not because sometimes the Human touch is and will continue to be needed.

Interviewer:

Yes, because that's exactly my next question then, because my next question was about human control.

Q2.4a: Of course, there are certain things that can be automated, but still need human control. And yes, some already indicated in the previous interviews that not everything still needs human control. How do you yourself feel about that?

Respondent:
I think ultimately that has to be risk based. And also especially impact based. LLMs are perfectly capable of indicating how confident they are in their response, especially if you use APIs. In the API you can see to what extent those vectors actually match, and if that percentage is very high and the potential impact of the wrong match is relatively small. Then I have no problem with applying that without human intervention. That example I gave just now of matching open invoices to incoming payments. Most, of course, in the customer base is just paid by direct debit, but a small portion still just writes to the moderate the premium of a policy. Well, then you have to include a policy number or a relationship number in there. You can imagine that there are still people there who put a typo in there or put nothing at all or whatever yes.
Then you have to look at match, the amount, what name is on the account, those kinds of aspects well, then you can ask yourself, how bad is it if for a premium of 3 tens and a probably a 99% match I just say okay, I consider that premium paid or maybe more annoyingly maybe not paid. Next thing you know there's a bailiff at the door. I think you always have to look at that from what the consequences are, because I think there are other examples In the world where those consequences are many times greater.

Interviewer:

No, great then I also had a question about the data, because what I had also noticed in my previous interviews, is that there was kind of a concern , kind of about synthetic data, so data that has already been generated by LLMs. And that it can occur that it can be used as input data.

Q2.5: How important do you consider the quality, provenance and control of the data used in AI applications such as LLMs?

Respondent:

I think it is very difficult to validate that properly, because how do you determine what is true? The Internet is also taking information from each other at lightning speed. You could say: Only when something occurs at least ten times do I consider it to be true, but even that is no guarantee.

At the same time, LLMs should get better and better at recognizing patterns in the data with which they are trained. A model should be able to distinguish whether a text comes from a human or another LLM. Personally, I do not find it necessarily problematic if LLMs are trained on output from other models, as long as that output is of high quality. The real risk is in the fact that you can ask an LLM to generate complete nonsense, and then that nonsense ends up in the training data. That's worrisome.

Interviewer:

I also had a question about the controls of employing an LLM:

Q2.6: In what ways does your organization control how LLMs are deployed, updated or modified? For example, think about prompt filtering etc.

Respondent:

I think you should indeed check on the prompt. I also think that you. would actually apply a certain filtering on the response if it is undesirable in terms of racism or other aspects.

That you don't let that response go through. Yes, because I think your prompt filtering can be very strong, but if you're smart enough, you can still formulate a prompt that gets through the prompt filtering that still writes per major nonsense or undesirable response.

I definitely also think you need to think about who is using what models. If you integrate everything into applications, what models do I want to be able to use?

Some models are much more prone to giving certain responses than others. So You can also make certain choices there at the time you integrate it into your application to just not make certain models available in that API.

Interviewer:
Well great, most of the key questions have already been covered and some you had already answered in another question. I would now suggest moving on to the framework itself.
Have you already read through the framework yourself?

Respondent:

Yes, I will get it to you.

Interviewer:
Well, a little introduction or the framework. So the framework is based on the literature analysis I did for this. So that consisted of governance and compliance laws. And what LLMs are at all. So from that I decided that the Orange boxes represent boxes where LLMs can actually be applied right away and the green is where it can at least help, but it doesn't necessarily have to be used. So I have 6 elements and then my first question is immediately:

Q2.7: What would you need to have confidence in an AI-supported framework for cybersecurity risk management?

Respondent:

To have confidence in such a framework, it is important to me that the various components are clear and understandable. Some terms, such as "data quality & information governance," can be open to interpretation. If you provide a clear explanation for that, it adds to the applicability and credibility of the whole thing. In addition, governance plays a key role. A framework can be technically sound, but if it is not clear who is responsible for what, or who supervises it, it will not work in practice. So for me, "Governance and Oversight" is the foundation on which everything else rests.
Without that, implementation just doesn't work.

Interviewer:

Q2.8: What elements do you find most relevant or problematic in your practice? Are you missing certain elements or do you see opportunities for improvement?

***Respondent*:**

Yes, where does this fall in for you is detecting deviations in the existing measures. So basically detecting incidents.

Interviewer:

So in the case of incidents I thought to myself that I could do that with the risk identification, but after the interview of before, it was recommended to make a separate element that could then make a connection with a risk identification.

Respondent:

Yes, I think so, too. Because, I think risks are really just something you come up with and incidents are actually the materialization of your risks.

Interviewer:

So that would go to Risk Identification. And also what you said so particularly here: education. So how can it be used? I also got a tip about incorporating that into the framework. I myself thought that could be added to AI Life Cycle.Do you have any further comment on that? It can basically be a separate element as well.

***Respondent*:**

Could also be a separate element, it could also be part of governance.

Interviewer:

And, why do you think that?

***Respondent*:**

Because I think training is part of your cognos model. But I think ultimately a separate block makes more sense.

Interviewer:
That's good to know. Those were basically my questions about the framework. Do you have any comments or recommendations of your own regarding the use of LLMs in cybersecurity? Or do you think we have already covered the main points?

Respondent:
No, I think we have covered most things already.

Closure:
Interviewer:
Okay, great. The idea is that I am going to incorporate the results of this interview into the framework, so it will be modified a little bit based on that. Then, in the second week of June, I will send a short Google survey to all the participants of the interviews. In doing so, I will also include the updated version of the framework so that I can validate it.

The survey will consist of about five to six simple, closed-ended questions, and I immediately give it a deadline. I am sending it out early in the second week of June, with the deadline being the end of that same week.

That brings us to the end of this interview. Thank you again for your time and valuable insights, I really got a lot out of this. As previously stated, all responses will be completely anonymized and deleted immediately upon completion of my thesis. Nothing will be stored or used for other purposes.

Respondent:

great, You're welcome and good luck!