Interview Guide (Read the notes at the bottom of the page)

1. Present ourselves.
2. Present the interview goals. (which includes we would like to hear your professional experiences and opinions for each question)
3. Ask for their Consent.
4. We will be storing the data provided by you for further studies apart from using them for our Master thesis and the documented report.
5. We will handle all the information you provide with confidentiality. Your name or the company's name will not be either published or used anywhere in the thesis report or any other places in relation to your answers.
6. Can we record?

Research Questions (RQs will be removed from Interview Guide - Added for easy reference here)

RQ1: What are the challenges encountered in practice when deriving requirements for AI components in particular concerning the selection of training data for safety-critical applications?

RQ2: What is the role of runtime monitoring in defining safety requirements and supporting safety argumentation?

Set of Questions that we want to ask (DRAFT) - OVERVIEW VERSION

Questions about the Role

1. Describe shortly your role and what you work with? Priority
2. What is your experience in this field of work? Priority
3. What does your team focus on? What kind of product/system is your team working on and what is the goal of the product/system? Priority

Questions on Training data

4. What is the typical " training data" that you and your team are dealing with? Priority
    a. Does the amount of data affect the behavior of the system and how?
5. How do you select the data to build your model to perform the required action of your system? Priority
6. What are the challenges while selecting the data? Priority
    a. Is it a challenge to obtain enough data ?
    b. Challenges in annotation ?
7. Have you written any set of requirements for selecting data? Priority
    a. What challenges do you encounter while setting requirements for selecting data?

b. **How would a good data specification look like? Priority**
c. What components would be part of a data specification? Priority

Questions on ML Components

8. What do you consider as a safety-critical system?
9. Are you working with any safety-critical systems? If needed, explain what a safety-critical system is. Priority
10. Are you familiar with the Safety Standards?
    a. What standards come to your mind .,(wait for a bit and ask) is it (ISO26262) or anything else?
    b. What would be the characteristics of the safety-critical systems?
    c. In what way do you think these Safety standards influence the system?
11. **Are you incorporating any ML into the safety-critical system? What do you think is the major challenge with incorporating ML into a safety-critical system? (If they don't use ML,) Why don't you use it in your safety-critical**

Questions on Run time Monitoring

12. Are you familiar with the runtime monitoring of the Machine learning part of the system? If not familiar, explain and ask, What do you think about monitoring the safety-critical during runtime? Priority
13. How to perform the run-time checks for the specification models? What do they envision from the run time checks? ? Priority
    a. Do you think runtime monitoring helps us in identifying the training data better? Priority
14. Explain the scenario, (to explain the bias that the system faces due to inadequate data) How does observing the data that the system is fed into during runtime help us in understanding the behavior of the system? Priority

Conclusion

Summarize briefly the most relevant points.

Is there something we should have asked about but did not?

Is there anything you would like to add? Or missing? Or mention?

Thank you! Follow-up questions via email?/Want an update on our thesis/report?

Time Calculations :

Current number of questions - 14

Approx time for each question - 4 (mean approximation)

Total app time - 14*4 = 56, if we keep 4.5 minutes each , total time is 63 (which is reasonable)

Total headings: 4 , 56/4 = 14 minutes each subsections, 63/4 = 15.75

We got to reduce the time in the first section and second section, focus more on 3, 4.

I would suggest Section 1- 6 minutes, Section 2- 17 minutes, Section 3- 20 minutes, Section 4- 17 minutes

Still its total of 60 minutes, Currently, we have 3 versions-

Version 0.1- First one we wrote and showed to ---------- and Hans,

Version 0.2 - The one we showed to Eric and Hans

Version 0.3- Changes based on 25th Feb meeting

Version 0.4- Current version