Q: Did you participate in any discussion about user interactions with the model?

A: Initially, we discussed what data we would capture from the user before discussing the model. There was no discussion with the whole team but partial discussions with business stakeholders, data scientists, among other members. At that stage, we were still defining product concepts, whose meanings differed for each actor.

We had several discussions with client representatives to understand their product vision and define what was and was not possible. From there, the UX designers started to prototype ideas we used to model the system database. Not all software engineers participated in these meetings, as having all of them participating could have affected our productivity.

Q: Did you notice any difficulties during the discussions with the data science team?

A: I did not experience any difficulties. Since the data scientists were part of a separate team, their activities were like a "black box" that we did not need to care about. We only asked them what data they needed, and then we included a new field in the UI so the user could fill it in. I know that internally they did a more detailed analysis of the data. But I was not a part of those discussions, as I was from another team.

Q: Did you participate in discussions about obtaining the data for the model?

A: We only had a few discussions about this because we mainly depended on how the client representatives could provide us with the data. So we just obtained the data following the methodology they defined and adapted it to how the data scientists desired. I do not know if this was the ideal way to get the data, but it was how we handled it in our scenario. Our documentation of this process consists of emails we could retrieve in case of an audit.

The data scientists defined what data was needed, but our team took full responsibility for obtaining the data. We discussed this subject with them to a certain extent, but since this process required skills they did not have, it was easier for us to assume this responsibility.

Q: Did you participate in any data-related activity?

A: We only captured the data needed to train the model.

Q: Did you participate in the definition of how the system would consume the model?

A: At the beginning of the project, we discussed this with the data science team. We created a REST API to allow the model integration with the system. We defined a communication interface for the API, and then each team did its part. It was outside the data science team's interest to understand how we stored the data as long as this service existed. So this responsibility was left to us, and we discussed it only among ourselves.

Q: How was the discussion about the storage of the model artifacts?

A: We did not have that discussion, which led to problems. We provided Git repositories for this storage, but the teams did not discuss how the data scientists would store the artifacts. This eventually caused issues because the model had a lot of artifacts, such as the training scripts, which were not separated from the API code. For this reason, large files were loaded unnecessarily every time a new model release was generated.

Q: Was any data scientist involved in the publication of the model?

A: No. We already had a pattern for deployment beforehand, and we knew the data scientists did not specialize in DevOps, so we left this structure ready for them. All the infrastructure concerning the model is our responsibility.

I also participated in the selection of machines for model training. The training infrastructure is heavy and demands good machines. I did not participate in the definitions of how the training would work.

Q: Was there any discussion about model updates and implementing incremental learning?

A: I did not participate in any discussion on this subject. I do not know what was decided.

Q: How was the discussion regarding model monitoring?

A: The responsibility for model performance lies with the data science team. The service created to consume the model is our responsibility.

Q: Were security and privacy issues considered when building the system?

A: Yes. The system's backend is the one that communicates with the model, which improves security. We had several discussions about this topic when designing the product, as it is available online.

Q: How was the relationship with the data scientists during the integration of the model with the rest of the system?

A: It was natural. We defined the data needed in the discussions between the teams. We then defined an API from there, and each team followed its side.

Still, we had some problems with changes in the communication interface established for the API, and also with the timing of these changes. But I think these issues are more related to managing changes in requirements and tasks than the teams' separation. If both teams were

closer, we could have avoided these problems. However, we would also have had a higher overhead, as everyone would need to be together in all meetings.

Q: Do you think actors linked to requirements management could have participated more?

A: Yes, this participation is very important, especially when you are delivering a product to a client. System definitions are expected to change throughout the project, so it is important to manage these changes effectively. The client representatives' expectations and what was passed on to the team must be clear.

Q: Is the integration of the model with the rest of the system well-documented?

A: Yes, I think so. Our biggest challenge was regarding the changes. The system's initial state was well-documented, but then changes started happening. These changes were not documented properly, which harmed the alignment between the teams. We did not correctly update the documentation throughout the project, and we also did not communicate these changes efficiently. We discovered them as system components stopped working.