**Second meeting with client (Florian G., at FlyBot’s Offices)**

**Discussed points:**

* Discussed the **“Evaluating NLU services for Conversational Question Answering Systems”** paper, which offers a state-of-the-art benchmark of conversational chatbot services. The  **RASA NLU** service emerged as the best choice: amongst the best metrics (not the best, the best one is **LUIS**), and free/open source
* The client gave an in-depth presentation of the tool they are currently using: **wit.ai**, and discussed the limitations of this tool, some problems they faced whilst using it, and the improvements they would like to see in our deliverables. The client agreed to share with the group the  **wit.ai flybot application**, which contains their entities (a total of 7), and offers the possibility to test the chatbot and see in real time the extracted entities with associated confidence metric
* The client shared with the group **the log file** which contains the data they collected through the use of flybot (some 17 million lines of entities). This file is likely to serve as a training basis for the developed tool. The client also suggested to generate more similar data, using an algorithm they developed, to have more data for training/validation
* The **specifications (cahier des charges) document** was discussed: a gdoc will be created and shared between all members of the group and the client to collaboratively elaborate the said document. So far, the main objectives are: Develop an open source chatbot, very performant in travel, which can eventually be trained to cover all tourism activities. To this end, a  **configuration file** should also be included in the deliverables which would enable the FlyBot team to easily add new entities, and thus expand the use cases of the chatbot. Moreover, an API Rest will be included to send test phrases to the chatbot and get the entities. Finally, a less important objective is to use some of the client’s data and measures to carry out a basic statistical study (what do people look for the most? When? Seasonal trends? When is the Conversion rate the highest? etc)
* In order to develop the API, the client agreed to share with us an example (a python file) to guide the team in this task
* The possibility to use OSIRIM (or another high performance calculator) was discussed, in order to carry out processing actions on the big amount of available data. Moreover, the client suggested to get in touch with the IRIT Research Team **Melodie**, who work in the field of NLP/NLU and who can thus offer some guidance
* **NEXT STEPS:**
* Finalize the specifications document
* Establish a costing and execution plan