# Lab 2: Basic Dialogue Management

LT2216: Dialogue Systems
Erik Kolterjahn Kjellberg

February 20, 2025

## Limitations and improvements

The appointment booking model I constructed is able to process different weekdays, times, and persons as well as different variations of "yes" and "no". It also re-raises the question if it did not understand what the user said or if the user was quiet. However, the model is not very flexible, and has several limitations.

### Grammar limitations

One limitation which the model had is that when asked for day and time, only one word utterances such as "Monday" and "12:30" were accepted, but not "on Tuesday" or "at 14 o'clock". This was improved by changing my grammar slightly such that if for example the utterance starts with "on", this part is stripped and the resulting string is processed. This allowed for slightly more flexible ways to specify the day and time of the meeting. However, many related features which one would probably want in a real-life application are still missing. As an example, it would be beneficial if the model could process an utterance with context-dependent meaning, such as "tomorrow", make an inference from this and the current day, and book a meeting the next day in the calendar.

Another grammar-related limitation is that the system so far only supports four persons. In reality, the persons would probably be stored in some database of customers, employees or similar, rather than in the script itself.

### State machine limitations

The state machine for the bookings is not very flexible. The user can only enter the person, day and possibly time in that order, and if the user wants to change something later, they need to start over again, and enter all of the details again. Ideally, the user should be able to change only one thing when the machine asks for confirmation. However, as this would require restructuring the whole flow of the state machine, I have not implemented it.

### ASR and TTS limitations

I encountered some problems related to the recognition of names. Upon adding my own name, which is spelled with a "k" unlike the corresponding English name, to the vocabulary, the ASR system was not able to understand neither the Swedish nor the English pronunciation of it. Instead, I had to say each letter separately. This is something which could most likely be fixed by fine-tuning the ASR model.