Chatbot - SI

Mini Project 1

Group n.º 6

- Daniel Gomes, Nmec: 93015
- Mário Silva, Nmec: 93430
- João Carvalho, Nmec: 89059
- Rui Fernandes, Nmec: 92952
- Gonçalo Passos, Nmec: 88864

Work Objectives

For this assignment, the main goals of our work are:

- The development of a Conversational agent (Chatbot).
- NLP should be used for some common sentences types.
- Ability to accumulate information/knowledge provided by interlocutors (i.e. learn from interaction) and produce answers to questions.
- For grammatically incorrect sentences, or sentences not supported by the system, react in a “seemingly intelligent” way

To do so, we recurred to the ***SpaCy*** Library, as we will demonstrate ahead.

Natural Language Processing (NLP)

NLP is a branch of AI concerned with giving computers the ability to understand text and spoken words in much the same way human beings can. However, NLP is a generally hard process:

- Highly ambiguous - not easy to program disambiguation
- Vague (the principle of minimal effort) - difficult to program the context and a priori knowledge
- Universal (domain independent) - hard to program general knowledge representation

Library Used - Spacy

SpaCy is a free and open-source library for Natural Language Processing (NLP) in Python with a lot of in-built capabilities.

- increasingly popular for processing and analyzing data in NLP.
- designed to process and “understand” large volumes of text
- can be used to build information extraction or natural language understanding

SpaCy is not an out-of-the-box chatbot engine!

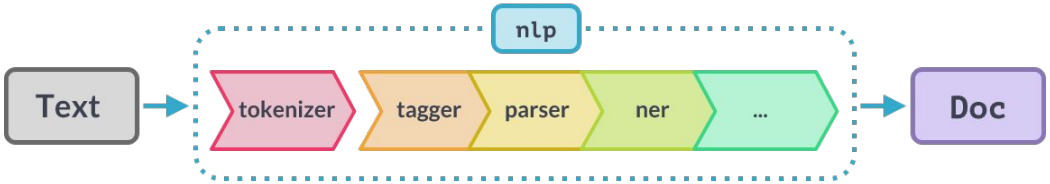
It provides features such as Tokenization, Lemmatization, Training or **Similarity**

How Spacy works?

The central data structures in spaCy are the **Language** class, the **Vocab** and the **Doc** object.

- **Language** class is used to process a text and turn it into a Doc object.
- **Doc** object owns the sequence of tokens and all their annotations.
- **Vocab** is a storage class for vocabulary and other data shared across a language

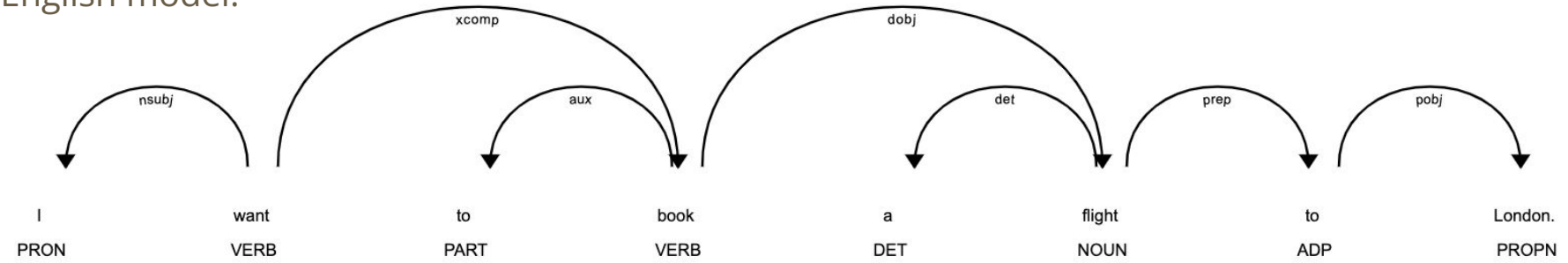
By centralizing strings, word vectors and lexical attributes in the Vocab, we avoid storing multiple copies of this data, this allows us to save memory.



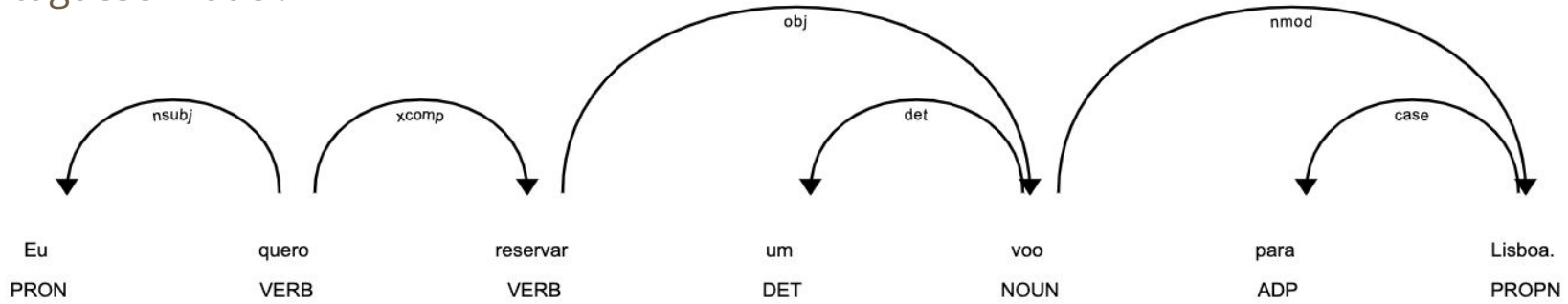
Spacy's Syntatic Structure Analysis of a Sentence

Spacy provides a several syntactic analysis methods, the following one performs the **Part of Speech Tagging**:

English model:



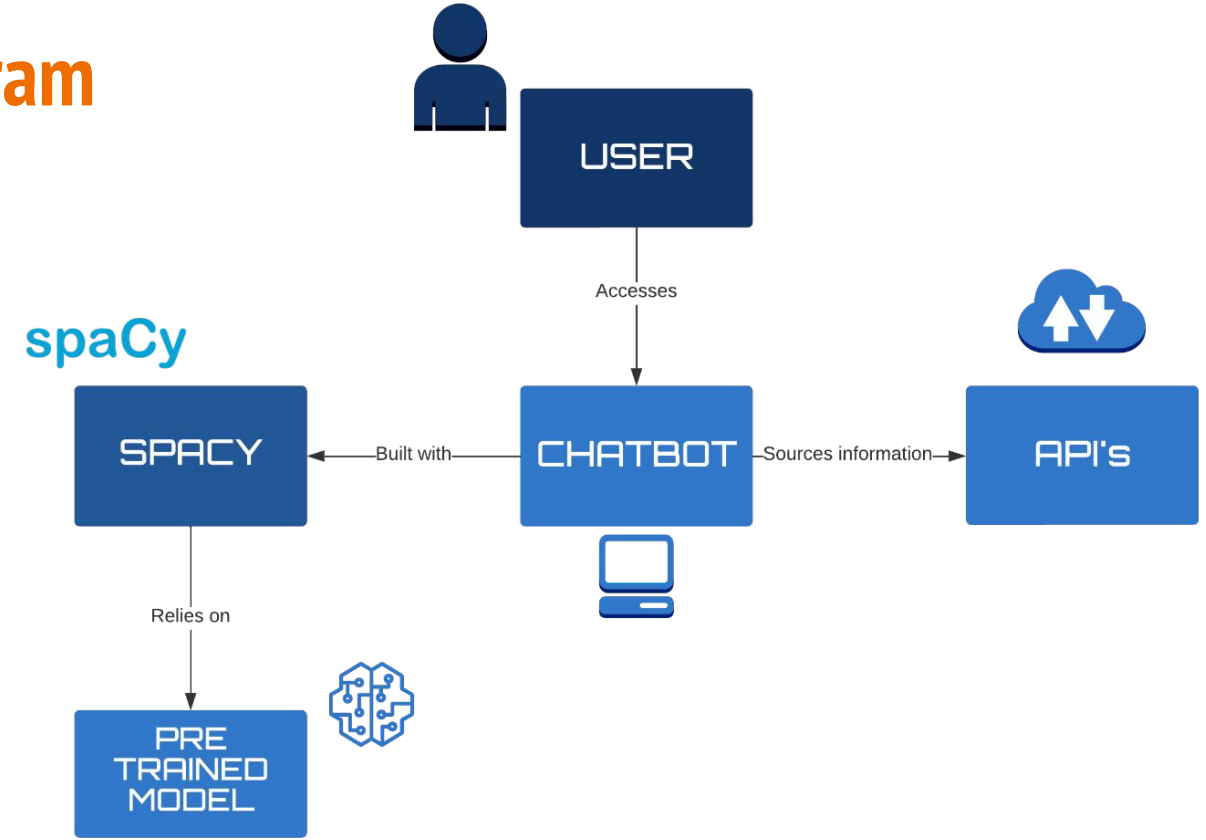
Portuguese model:



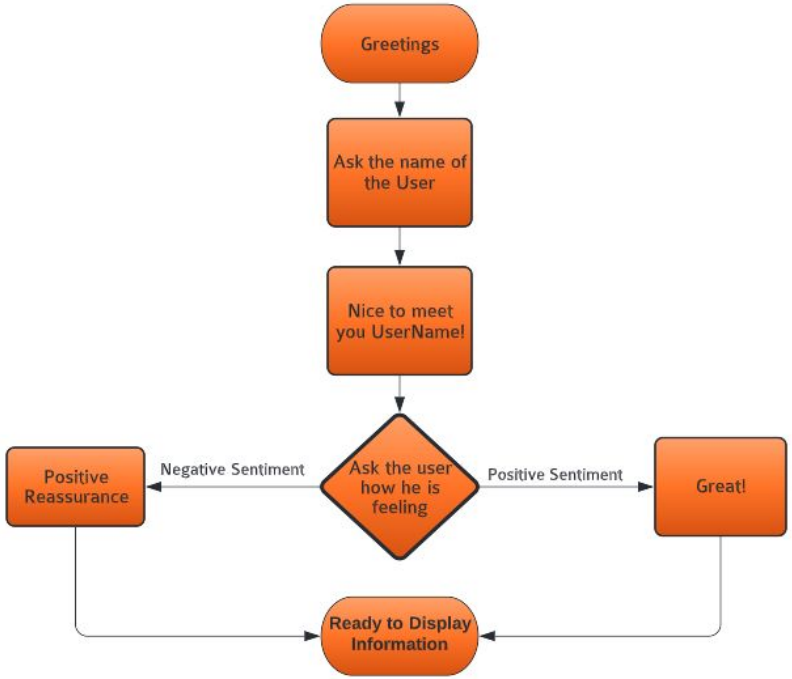
What is our Chatbot about?

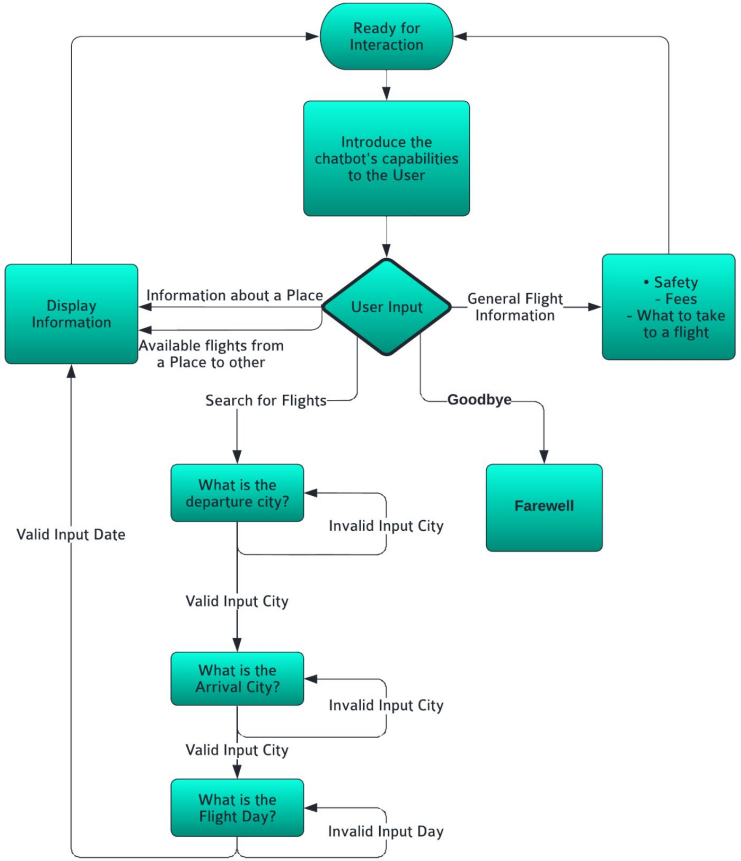
- Our chatbot is a friendly and informative AI system, that can seemingly communicate with users.
- Its main purpose is to provide important information about flight travelling, and allow searching for flights given departure and arrival cities.
- It also provides the ability to show information about desired places.
- Finally, it also collects feedback from the users, so we may keep track of the main downsides of the system.

Architecture Diagram

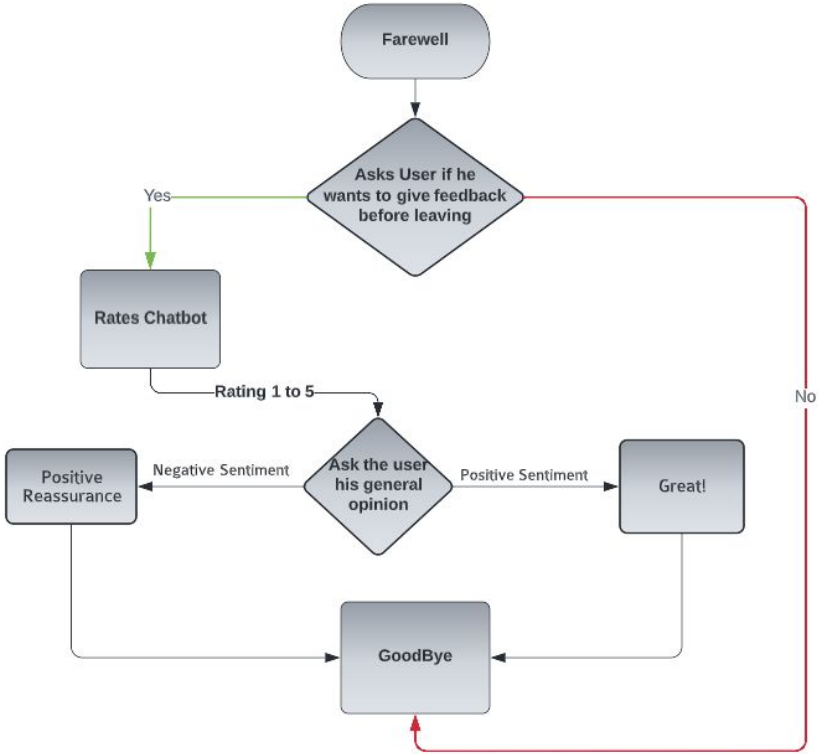


Flow Diagram





Flow Diagram



Demo