# Chatbot design planning document

-Elliot Harding (N0688119)

**Chosen topic area**

This project is the attempt at creating a dog breed information chatbot. The chatbot will attempt to provide its user with information on different dog breeds in a conversation like manner.

**List of requirements**

In order to achieve the aim of the project, a list of requirements must be met. For its first iteration the bot will be able to;

1. Perform conversation pleasantries.
2. Provide the user with information on requested dog breeds.
3. Provide the user on requested lists of groups of dog breeds (i.e. hounds, terriers, retrievers, etc…).
4. Provide user information on dog sizes.
5. Tell the user if a dog is a cross or pure breed.
6. Attempt to guess and provide information on the dog breed a user describes.
7. Tell the user dog related question-answer jokes.
8. Perform Wikipedia searches when it cannot provide requested information.

**Description of individual modules**

In order to achieve the requirements above, the chatbot implementation will utilize various modules.

TF-IDF

This module utilizes the term frequency, inverse document frequency and cosine similarity equations in order to calculate the similarity of a sentence in comparison to a list of other sentences. This is to be used by the chatbot to compare users input with its data to understand the users input, and therefore give a better response. The implementation of requirements 2 to 7 will use this module.

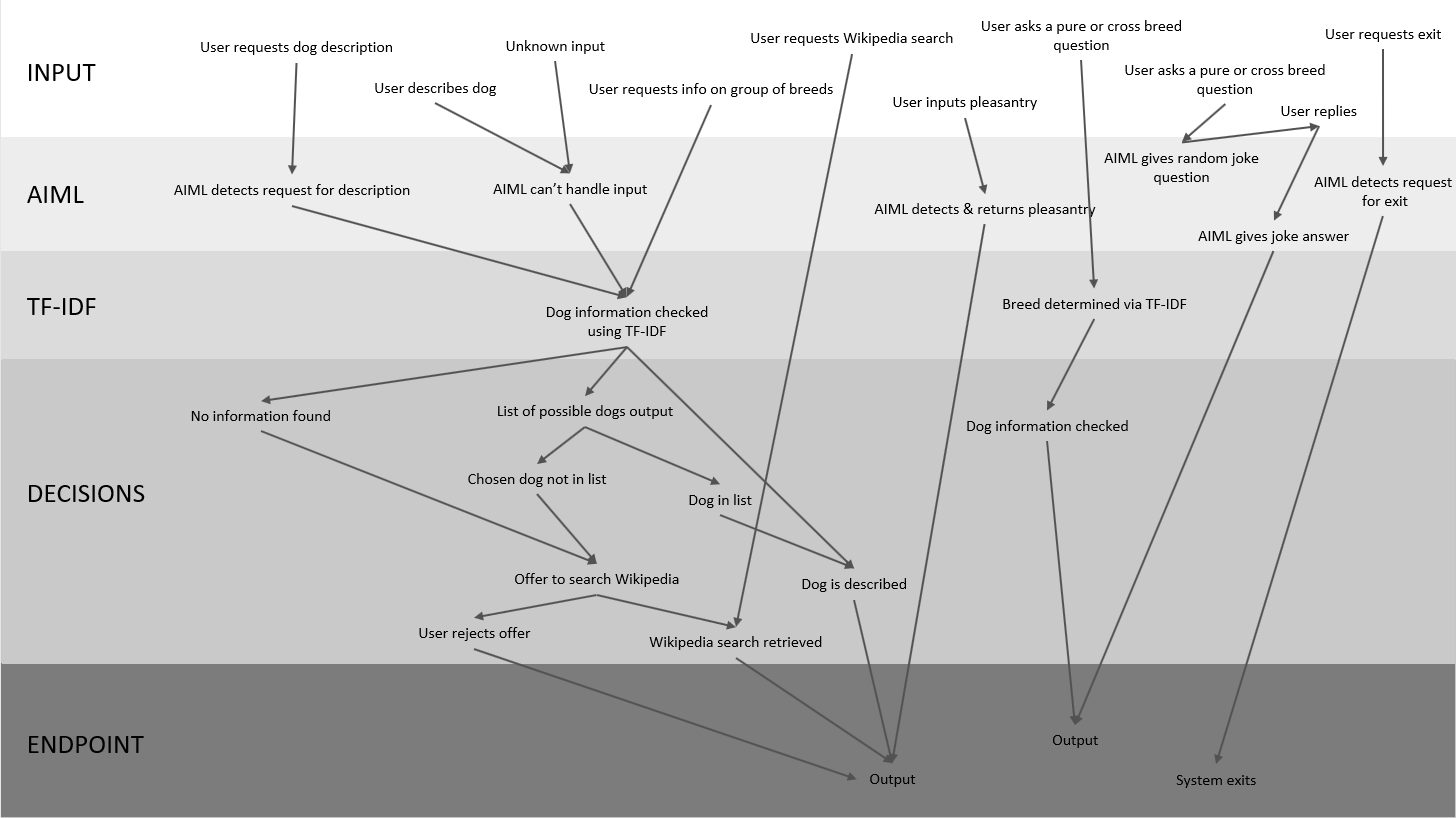
AIML

The AIML module will be used by the chatbot to define the response rules to users input. For example, to return information about a specific dog breed when users request it. Each input run of the chatbot will be initially handled by the AIML. It will decide the next steps of the chatbot algorithm in order to produce the correct output, expect however in the case of pleasantries and jokes in which their entire functionality will be handled by it.

Dog breed and information

Data pertaining to a breed’s name with its corresponding information will be stored within an CSV, like a 2D array. The chatbot functionality will use this data in order to achieve requirements 2 to 7.

**Flow diagram**

Below is a flow diagram representing the functionality flow between the various modules described above.

Mention the fact you're loading the tokenizer from a saved version

mention that wiki thingy is still chill cuz wiki command